

\*\*\*\*\*  
 W P E R L H  
 \*\*\*\*\*  
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998, University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jan 7 11:05:22 1999; MasPar time 7.06 Seconds  
 r output not generated. 592.942 Million cell updates/sec

Title: >US-08-799-910-10  
 Description: (1-156) from US08799910.pep  
 Perfect Score: 1114  
 Sequence: 1 MCHSRCHPTWILQAPTA.....EPSTDYALDLSTFLOQHPAAF 156

Scoring table: PAM 150  
 Gap 11

Searched: 74019 seqs, 26840295 residues

Post-processing: Minimum Match 0%  
 Listing first 200 summaries

Database: swiss-prot35  
 1:swissprot

Statistics: Mean 44.450; Variance 91.313; scale 0.487

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1114	100.0	156	1	1EX1_HUMAN	RADIATION-INDUCIBLE IM	3.79e-194
2	752	67.5	153	1EX1_MOUSE	RADIATION-INDUCIBLE IM	3.93e-121
3	117	10.5	474	1VTP3_TTVIV	VIRAL PROTEIN TPX.	2.12e-03
4	112	10.1	228	1VCOM_ADEM1	MINOR CORE PROTEIN (PR	1.01e-02
5	109	9.8	1799	1Y025_CAEEL	HYPOTHETICAL 202.6 KD	2.51e-02
6	103	9.2	1206	1FOR4_MOUSE	FORMIN 4 (LIMB DEFORMI	1.50e-01
7	103	9.2	1468	1FORM_MOUSE	FORMIN (LIMB DEFORMITY	1.50e-01
8	101	9.1	268	1CEBD_RAT	CCAAT/ENHANCER BINDING	2.68e-01
9	101	9.1	417	1AIAB_CANFA	ALPHA-1B ADRENERGIC RE	2.68e-01
10	101	9.1	816	1QAIF_NEUCR	QUINIC ACID UTILIZATION	2.68e-01
11	99	8.9	272	1GSPC_ERMCH	GENERAL SECRETION PATH	4.76e-01
12	98	8.8	272	1GSPC_ERMCH	GENERAL SECRETION PATH	4.76e-01
13	98	8.8	390	1VGLI_HSV11	GLYCOPROTEIN 1.	6.33e-01
14	98	8.8	475	1GAPN_STRMU	NADP-DEPENDENT GLYCERA	6.33e-01
15	98	8.8	519	1AIAB_HUMAN	ALPHA-1B ADRENERGIC RE	6.33e-01
16	98	8.8	975	1CDP_CANFA	CCAAT DISPLACEMENT PRO	6.33e-01
17	97	8.7	1043	1CHS2_PABER	CHITIN SYNTHASE 2 (EC	8.39e-01
18	97	8.7	1239	1V120_EBV	CAPSID ASSEMBLY PROTEI	8.39e-01
19	97	8.7	1337	1PTPJ_HUMAN	PROTEIN-TYROSINE PHOSP	8.39e-01
20	96	8.6	245	1V14_RPV5B	PROBABLE E4 PROTEIN.	1.11e+00
21	96	8.6	245	1V14_RPV5B	PROBABLE E4 PROTEIN.	1.11e+00
22	96	8.6	1233	1NME3_HUMAN	GLUTAMATE (NMDA) RECEP	1.11e+00
23	95	8.5	234	1GLNA_DUNSA	GLUTAMINE SYNTHETASE (	1.47e+00

1	CDNC_HUMAN	CYCLIN-DEPENDENT KINAS	1.47e+00
1	APG_BRANA	ANTER-SPECIFIC PROLINE	1.47e+00
1	GUNA_XANCP	MAJOR EXTRACELLULAR EN	1.47e+00
1	APG_ARATH	ANTER-SPECIFIC PROLINE	1.47e+00
1	VP40_HSV2	CAPSID PROTEIN P40 (CO	1.47e+00
1	HEX_MOUSE	ZINC FINGER PROTEIN HR	1.47e+00
1	YXR4_EBV	HYPOTHETICAL BKRF4 PRO	1.93e+00
1	Y091_NPVOV	HYPOTHETICAL 29.3 KD P	1.93e+00
1	NDPP_MOUSE	NPC DERIVED PROLINE RI	1.93e+00
1	GUNA_CALSA	ENDOGLUCANASE A PRECUR	1.93e+00
1	SAX1_CHICK	HOMEOBOX PROTEIN SAX-1	2.54e+00
1	CEBD_MOUSE	CCAAT/ENHANCER BINDING	2.54e+00
1	YIT4_YEAST	HYPOTHETICAL 31.3 KD P	2.54e+00
1	Y08N_MYCTU	HYPOTHETICAL 37.0 KD P	2.54e+00
1	SP62_HUMAN	SPICEOSOME ASSOCIATED	2.54e+00
1	MBEA_ECOLI	MOBILIZATION PROTEIN M	2.54e+00
1	EXTN_TOBAC	EXTENSIN PRECURSOR (CE	3.34e+00
1	DMK_MOUSE	MYOTONIN-PROTEIN KINAS	3.34e+00
1	FIBL_HUMAN	FIBRINOGEN-LIKE PROTEI	3.34e+00
1	YIK9_YEAST	HYPOTHETICAL 103.6 KD	3.34e+00
1	POLN_HEVPA	NON-STRUCTURAL POLYPRO	3.34e+00
1	ERMA_ARTS3	RNA ADENINE N-6-METHY	4.37e+00
1	YV19_MYCTU	HYPOTHETICAL 45.3 KD T	4.37e+00
1	SP62_MOUSE	SPICEOSOME ASSOCIATED	4.37e+00
1	SSP5_STRGN	AGGLUTININ RECEPTOR PR	4.37e+00
1	FAS2_YEAST	FATTY ACID SYNTHASE, S	4.37e+00
1	FGF4_CHICK	FIBROBLAST GROWTH FACT	5.71e+00
1	EXTN_SORVU	EXTENSIN PRECURSOR (PR	5.71e+00
1	YNS2_CAEEL	HYPOTHETICAL 43.2 KD P	5.71e+00
1	SECD_MYCTU	PROTEIN-EXPORT MEMBRAN	5.71e+00
1	GPBA_HUMAN	PLATELET GLYCOPROTEIN	5.71e+00
1	ABFA_STRLI	ALPHA-L-ARABINOFURANOS	5.71e+00
1	BAT2_HUMAN	LARGE PROLINE-RICH PRO	5.71e+00
1	YML2_THIFE	HYPOTHETICAL 12.3 KD P	7.44e+00
1	CEBD_HUMAN	CCAAT/ENHANCER BINDING	7.44e+00
1	VIEN_NPVAC	IMMEDIATE-EARLY REGULA	7.44e+00
1	DJNM_MYCPN	DNAJ-LIKE PROTEIN MG20	7.44e+00
1	LIO_DROME	LINOTTE PROTEIN.	7.44e+00
1	EBN4_EBV	EBNA-4 NUCLEAR PROTEIN	7.44e+00
1	LHR_ECOLI	PROBABLE ATP-DEPENDENT	7.44e+00
1	FRA2_HUMAN	FOS-RELATED ANTIGEN 2.	9.67e+00
1	IP35_STRFR	HYPOTHETICAL 35.5 KD P	9.67e+00
1	IP38_RAT	ID-MYO-INOSITOL-TRISPH	9.67e+00
1	VL96_IRV1	L96 PROTEIN.	9.67e+00
1	PH87_YEAST	INORGANIC PHOSPHATE TR	9.67e+00
1	DRPL_RAT	ATROPHIN-1 (DENTATORUB	9.67e+00
1	VPRT_HTLV2	PROTEIN RICH 28 KD ANT	1.25e+01
1	PR28_MYCTU	FOS-RELATED ANTIGEN 2.	1.25e+01
1	FRA2_RAT	GLYCOPROTEIN 1.	1.25e+01
1	VGLI_HSV23	GAG POLYPROTEIN (CONTA	1.25e+01
1	EBN2_EBV	EBNA-2 NUCLEAR PROTEIN	1.25e+01
1	PGH1_HUMAN	PROSTAGLANDIN G/H SYNT	1.25e+01
1	ERK5_HUMAN	EXTRACELLULAR SIGNAL-R	1.25e+01
1	DSC1_MOUSE	DESMOCELLIN 1A/1B PREC	1.25e+01
1	BOSS_DROME	BRIDE OF SEVENLESS PRO	1.25e+01
1	ICP4_HSV11	TRANS-ACTING TRANSCRIP	1.25e+01
1	CAL12_MOUSE	PROCOLLAGEN ALPHA 1(I	1.25e+01
1	POLN_HEVBU	NON-STRUCTURAL POLYPRO	1.25e+01
1	DYHC_EHNI	DYNEIN HEAVY CHAIN, CY	1.25e+01
1	CD52_MACFA	CAMPATH-1 ANTIGEN PREC	1.25e+01
1	I3MS_BRACA	MICROSPOR-SPECIFIC PR	1.62e+01
1	YH25_AZOC	HYPOTHETICAL 27.0 KD P	1.62e+01
1	REFX_HUMAN	REGULATORY FACTOR X-AS	1.62e+01
1	E1A_ADE02	EARLY E1A 32 KD PROTEI	1.62e+01
1	GLTC_WHEAT	GLUTENIN, LOW MOLECULA	1.62e+01
1	GAG_FSVHZ	GAG POLYPROTEIN (CONTA	1.62e+01
1	VCOM_ADECC	MINOR CORE PROTEIN (PR	1.62e+01
1	WNTC_DROME	PROTEIN DINT-1 PRECURS	1.62e+01
1	PDI_MEDSA	PROTEIN DISULFIDE ISOM	1.62e+01
1	BRL1_EBV	TRANSCRIPTION ACTIVATO	1.62e+01
1	NIA2_HORVU	NITRATE REDUCTASE (EC	1.62e+01



\*\*\*\*\*

W P S R L A

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998, University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Thu Jan 7 11:05:22 1999; MasPar time 7.06 Seconds  
r output not generated. 592.942 Million cell updates/sec

Title: >US-08-799-910-10  
Description: (1-156) from US08799910.pep  
Perfect Score: 1114  
Sequence: 1 MCHSRCHPTMTILOAPTA.....EPDVALDSTFLOQHPAAAF 156

Scoring table:

PAM 150  
Gap 11

Searched: 74019 seqs, 26840295 residues

Post-processing: Minimum Match 0%  
Listing first 200 summaries

Database: swiss-prot35  
1:swissprot

Statistics: Mean 44.450; Variance 91.313; scale 0.487

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1114	100.0	156	1	IEX1_HUMAN	RADIATION-INDUCIBLE IM	3.79e-194
752	67.5	153	1	IEX1_MOUSE	RADIATION-INDUCIBLE IM	3.93e-121
3	117	10.5	474	VTP3_TTV1V	VIRAL PROTEIN TPX.	2.12e-03
4	112	10.1	228	YCON_ADEMI	MINOR CORE PROTEIN (PR	1.01e-02
5	109	9.8	1799	V035_CAEEL	HYPOTHETICAL 202.6 KD	2.51e-02
6	103	9.2	1206	FORA_MOUSE	FORMIN 4 (LIMB DEFORMI	1.50e-01
7	103	9.2	1468	FORM_MOUSE	FORMIN (LIMB DEFORMIT	1.50e-01
8	101	9.1	268	CEBD_RAT	CCAAT/ENHANCER BINDING	2.68e-01
9	101	9.1	417	ALAB_CANFA	ALPHA-1B ADRENERGIC RE	2.68e-01
10	101	9.1	816	QAIF_NEUCR	QUINIC ACID UTILIZATIO	2.68e-01
11	99	8.9	272	GSQC_ERMCH	GENERAL SECRETION PATH	4.76e-01
12	98	8.8	272	GSQC_ERMCH	GENERAL SECRETION PATH	6.33e-01
13	98	8.8	390	VGLI_HSV11	GLYCOPROTEIN I.	6.33e-01
14	98	8.8	475	GAPN_STRMO	NADP-DEPENDENT GLYCERA	6.33e-01
15	98	8.8	519	ALAB_HUMAN	ALPHA-1B ADRENERGIC RE	6.33e-01
16	98	8.8	975	CDP_CANFA	CCAAT DISPLACEMENT PR	6.33e-01
17	97	8.7	1043	CHS2_PABRB	CHITIN SYNTHASE 2 (EC	8.39e-01
18	97	8.7	1239	V120_EBV	CAPSID ASSEMBLY PROTEI	8.39e-01
19	97	8.7	1337	PTPO_HUMAN	PROTEIN-TYROSINE PHOSP	8.39e-01
20	96	8.6	245	VE4_HPV5B	PROBABLE E4 PROTEIN.	1.11e+00
21	96	8.6	245	VE4_HPV05	PROBABLE E4 PROTEIN.	1.11e+00
22	96	8.6	1233	NME3_HUMAN	GLUTAMATE (NMDA) RECEP	1.11e+00
23	95	8.5	234	GLNA_DUNSA	GLUTAMINE SYNTHETASE (	1.47e+00

24	95	8.5	316	CDNC_HUMAN	CYCLIN-DEPENDENT KINAS	1.47e+00
25	95	8.5	449	APG_BRANA	ANTER-SPECIFIC PROLINE	1.47e+00
26	95	8.5	493	GUNA_XANCP	MAJOR EXTRACELLULAR EN	1.47e+00
27	95	8.5	534	APG_ARATH	ANTER-SPECIFIC PROLINE	1.47e+00
28	95	8.5	643	VP40_HSV2	CAPSID PROTEIN P40 (CO	1.47e+00
29	95	8.5	3866	ZINC_FINGER	ZINC FINGER PROTEIN HR	1.47e+00
30	94	8.4	217	YKR4_EBV	HYPOTHETICAL BKR4 PRO	1.93e+00
31	94	8.4	279	NDPP_MOUSE	HYPOTHETICAL 29.3 KD P	1.93e+00
32	94	8.4	389	GUNA_CALSA	ENDOGLUCANASE A PRECUR	1.93e+00
33	94	8.4	1742	SAXI_CHICK	HOMEOBOX PROTEIN SAX-1	2.54e+00
34	93	8.3	232	CEBD_MOUSE	CCAAT/ENHANCER BINDING	2.54e+00
35	93	8.3	268	YIT4_YEAST	HYPOTHETICAL 37.3 KD P	2.54e+00
36	93	8.3	276	Y08N_MYCTU	HYPOTHETICAL 31.0 KD P	2.54e+00
37	93	8.3	350	SP62_HUMAN	SPICOSOME ASSOCIATED	2.54e+00
38	93	8.3	464	MBEA_ECOLI	MOBILIZATION PROTEIN M	2.54e+00
39	93	8.3	517	EXTN_TOBAC	EXTENSIN PRECURSOR (CE	3.34e+00
40	92	8.3	620	YNS2_CAEEL	SPICOSOME ASSOCIATED	4.37e+00
41	92	8.3	631	DMK_MOUSE	MYOTONIN-PROTEIN KINAS	3.34e+00
42	92	8.3	820	FIBL_HUMAN	FIBRINOGEN-LIKE PROTEI	3.34e+00
43	92	8.3	926	YIK9_YEAST	HYPOTHETICAL 103.6 KD	3.34e+00
44	92	8.3	1692	POLN_HEVPA	NON-STRUCTURAL POLYPRO	3.34e+00
45	91	8.2	340	ERMA_ARPS3	RNA ADENINE N-6-METHY	4.37e+00
46	91	8.2	440	YW19_MYCTU	HYPOTHETICAL 45.3 KD T	4.37e+00
47	91	8.2	485	SP62_MOUSE	SPICOSOME ASSOCIATED	4.37e+00
48	91	8.2	1500	SSP5_STRGN	AGGLUTININ RECEPTOR PR	4.37e+00
49	91	8.2	1894	FAS2_YEAST	FATTY ACID SYNTHASE, S	4.37e+00
50	90	8.1	194	FGF4_CHICK	FIBROBLAST GROWTH FACT	5.71e+00
51	90	8.1	283	EXTN_SORVU	EXTENSIN PRECURSOR (PR	5.71e+00
52	90	8.1	455	YNS2_CAEEL	HYPOTHETICAL 43.2 KD P	5.71e+00
53	90	8.1	573	SECD_MYCTU	PROTEIN-EXPORT MEMBRAN	5.71e+00
54	90	8.1	626	GPBA_HUMAN	PLATELET GLYCOPROTEIN	5.71e+00
55	90	8.1	662	ABFA_STRLI	ALPHA-L-ARABINOFURANOS	5.71e+00
56	90	8.1	2142	BAT2_HUMAN	LARGE PROLINE-RICH PRO	5.71e+00
57	89	8.0	108	YWL2_THIFE	HYPOTHETICAL 12.3 KD P	7.44e+00
58	89	8.0	269	CEBD_HUMAN	CCAAT/ENHANCER BINDING	7.44e+00
59	89	8.0	408	VIEN_NPVAC	IMMEDIATE-EARLY REGULA	7.44e+00
60	89	8.0	910	DNJM_MYCPN	DNAJ-LIKE PROTEIN MG20	7.44e+00
61	89	8.0	915	LIO_DROME	LINOTTE PROTEIN.	7.44e+00
62	89	8.0	938	EBN4_EBV	EBNA-4 NUCLEAR PROTEIN	7.44e+00
63	89	8.0	1538	LHR_ECOLI	PROBABLE ATP-DEPENDENT	7.44e+00
64	88	7.9	326	FR22_HUMAN	FOS-RELATED ANTIGEN 2	9.67e+00
65	88	7.9	348	YI35_STRFR	HYPOTHETICAL 35.5 KD P	9.67e+00
66	88	7.9	459	IP3K_RAT	ID-MYO-INOSITOL-TRISPH	9.67e+00
67	88	7.9	867	VL96_IRV1	L96 PROTEIN.	9.67e+00
68	88	7.9	923	PH87_YEAST	INORGANIC PHOSPHATE TR	9.67e+00
69	88	7.9	1183	DEPL_RAT	ATROPHIN-1 (DENTATORUB	9.67e+00
70	87	7.8	178	VPRT_HTLV2	PROTEASE (EC 3.4.23.-)	1.25e+01
71	87	7.8	294	SMN1_HUMAN	SURVIVAL MOTOR NEURON	1.25e+01
72	87	7.8	310	PR28_MYCTU	PROLINE RICH 28 KD ANT	1.25e+01
73	87	7.8	327	FR22_RAT	FOS-RELATED ANTIGEN 2	1.25e+01
74	87	7.8	372	VGLI_HSV23	GLYCOPROTEIN I.	1.25e+01
75	87	7.8	425	GAG_ESVGA	GAG POLYPROTEIN (CONTA	1.25e+01
76	87	7.8	487	EBN2_EBV	EBNA-2 NUCLEAR PROTEIN	1.25e+01
77	87	7.8	599	PGH1_HUMAN	PROTAGLANDIN G/H SYNT	1.25e+01
78	87	7.8	815	ERK5_HUMAN	EXTRACELLULAR SIGNAL-R	1.25e+01
79	87	7.8	886	DSC1_MOUSE	DESMOCOLLIN 1A/1B PREC	1.25e+01
80	87	7.8	896	BOSS_DROME	BRIDE OF SEVENLESS PRO	1.25e+01
81	87	7.8	1298	ICP4_HSV11	TRANS-ACTING TRANSCRIPT	1.25e+01
82	87	7.8	1459	CA12_MOUSE	PROCOLLAGEN ALPHA 1(I)	1.25e+01
83	87	7.8	1693	POLN_HEVBU	NON-STRUCTURAL POLYPRO	1.25e+01
84	87	7.8	4344	DYHC_EMENI	DYNEIN HEAVY CHAIN, CY	1.62e+01
85	86	7.7	60	CD52_MACFA	CAMPATH-1 ANTIGEN PREC	1.62e+01
86	86	7.7	156	I3M5_BRANA	MICROSPORE-SPECIFIC PR	1.62e+01
87	86	7.7	245	YH25_AZOC	HYPOTHETICAL 27.0 KD P	1.62e+01
88	86	7.7	272	REXA_HUMAN	REGULATORY FACTOR X-AS	1.62e+01
89	86	7.7	289	EIA_ADE02	EARLY EIA 32 KD PROTEI	1.62e+01
90	86	7.7	295	GLTC_WHEAT	GLUTENIN, LOW MOLECULA	1.62e+01
91	86	7.7	414	GAG_FSVH2	GAG POLYPROTEIN (CONTA	1.62e+01
92	86	7.7	421	VCOM_ADECC	MINOR CORE PROTEIN (PR	1.62e+01
93	86	7.7	468	WNTG_DROME	PROTEIN DINT-1 PRECURS	1.62e+01
94	86	7.7	512	PDI_MEDSA	PROTEIN DISULFIDE ISOM	1.62e+01
95	86	7.7	605	BRL1_EBV	TRANSCRIPTION ACTIVATO	1.62e+01
96	86	7.7	912	NIA2_HORVU	NITRATE REDUCTASE (EC	1.62e+01

97	86	7.7	915	1	N1A1_HORVU	NITRATE REDUCTASE (EC	1.52e+01	170	82	7.4	282	1	HES1_MOUSE	TRANSCRIPTION FACTOR H	4.44e+01
98	86	7.7	1238	1	PTPJ_MOUSE	PROTEIN-TYROSINE PHOSP	2.10e+01	171	82	7.4	289	1	YFK2_YEAST	HYPOTHETICAL 31.9 KD P	4.44e+01
99	85	7.6	166	1	VP06_BPRPD	PROTEIN P6.	2.10e+01	172	82	7.4	353	1	CCPA_ACEXY	CELLULOSE COMPLEMENTIN	4.44e+01
100	85	7.6	190	1	BC77_SHEEP	BACTENECIN 7 PRECURSOR	2.10e+01	173	82	7.4	360	1	VTPX_TTV1	VIRAL PROTEIN TFX.	4.44e+01
101	85	7.6	205	1	PSAD_HORVU	PHOTOSYSTEM I REACTION	2.10e+01	174	82	7.4	368	1	MYCL_MOUSE	L-MYC PROTO-ONCOGENE P	4.44e+01
102	85	7.6	207	1	PLCR_PSEAE	PHOSPHOLIPASE C ACCESS	2.10e+01	175	82	7.4	379	1	CYBA_STELO	CYTOCHROME B (EC 1.10.	4.44e+01
103	85	7.6	236	1	CLAT_MCMV	COAT PROTEIN (CAPSID P	2.10e+01	176	82	7.4	388	1	VE2_HPV27	REGULATORY PROTEIN E2.	4.44e+01
104	85	7.6	258	1	UL49_HSVBP	TEGUMENT PROTEIN UL49	2.10e+01	177	82	7.4	442	1	ODO2_RAT	DIHYDROLIPOAMIDE SUCCI	4.44e+01
105	85	7.6	308	1	YL58_CAEEL	HYPOTHETICAL 34.2 KD P	2.10e+01	178	82	7.4	453	1	CARA_NEUCR	CARBAMOYL-PHOSPHATE SY	4.44e+01
106	85	7.6	379	1	CYB_STEAT	CYTOCHROME B (EC 1.10.	2.10e+01	179	82	7.4	472	1	VL2_HPV58	MINOR CAPSID PROTEIN L	4.44e+01
107	85	7.6	379	1	D4BR_MOUSE	CYTOCHROME B (EC 1.10.	2.10e+01	180	82	7.4	514	1	RP54_AZOCA	RNA POLYMERASE SIGMA-5	4.44e+01
108	85	7.6	387	1	CYBR_MOUSE	D(4) DOPAMINE RECEPTOR	2.10e+01	181	82	7.4	524	1	VL2_HPV22	MINOR CAPSID PROTEIN L	4.44e+01
109	85	7.6	506	1	VL2_HPV50	MINOR CAPSID PROTEIN L	2.10e+01	182	82	7.4	540	1	HUP2_CHLKE	H(+)/HEXOSE COTRANSFOR	4.44e+01
110	85	7.6	513	1	CK15_HUMAN	VOLTAGE-GATED POTASSIU	2.10e+01	183	82	7.4	575	1	VGLE_HSVBS	GLYCOPROTEIN E PRECURS	4.44e+01
111	85	7.6	755	1	RM1_HUMAN	RAS-RESPONSIVE ELEMENT	2.10e+01	184	82	7.4	589	1	VF40_SCMVC	CAPSID PROTEIN P40 (CO	4.44e+01
112	85	7.6	878	1	IL3B_MOUSE	INTERLEUKIN-3 RECEPTOR	2.10e+01	185	82	7.4	590	1	VG28_HSV11	HYPOTHETICAL GENE 28 P	4.44e+01
113	85	7.6	893	1	BOSS_DROVI	BRIDGE OF SEVENLESS PRO	2.10e+01	186	82	7.4	689	1	CN4D_HUMAN	CAMP-DEPENDENT 3',5'-C	4.44e+01
114	85	7.6	960	1	LI36_CAEEL	LIN-36 PROTEIN	2.10e+01	187	82	7.4	797	1	PAT1_YEAST	TOPOISOMERASE II-ASSOC	4.44e+01
115	85	7.6	1167	1	SCPA_STRPY	C5A PEPTIDASE PRECURSO	2.10e+01	188	82	7.4	839	1	V2A_CMVQ	2A PROTEIN (PROBABLE R	4.44e+01
116	85	7.6	1213	1	FORM_CHICK	FORMIN (LIMB DEFORMITY	2.10e+01	189	82	7.4	896	1	CYRB_MOUSE	CYTOKINE RECEPTOR COM	4.44e+01
117	85	7.6	1331	1	MANB_CALSA	BETA-MANNANASE / ENDOG	2.10e+01	190	82	7.4	1185	1	DRPL_HUMAN	ATROPHIN-1 (DENTATORUB	4.44e+01
118	85	7.6	1804	1	YF47_YEAST	HYPOTHETICAL 207.6 KD	2.10e+01	191	82	7.4	1446	1	IE18_PPVKA	IMMEDIATE-EARLY PROTEI	4.44e+01
119	85	7.6	1844	1	POLR_TYMV	RNA REPLICASE POLYPROT	2.10e+01	192	82	7.4	1465	1	DPOA_MOUSE	DNA POLYMERASE ALPHA (	4.44e+01
120	85	7.6	1844	1	POLR_TYMV	RNA REPLICASE POLYPROT	2.10e+01	193	82	7.4	1478	1	BCK1_YEAST	SERINE/THREONINE PROTE	4.44e+01
121	85	7.6	2175	1	HMCU_DROME	HOMEOBOX PROTEIN CUT.	2.10e+01	194	82	7.4	1487	1	ICP4_HSVEB	TRANS-ACTING TRANSCRIP	4.44e+01
122	83	7.5	293	1	H5_ANGAN	HISTONE H5.	3.46e+01	195	82	7.4	1487	1	ICP4_HSVBK	TRANS-ACTING TRANSCRIP	4.44e+01
123	83	7.5	293	1	VP10_BPRPD	PROTEIN P10.	3.46e+01	196	82	7.4	1487	1	ICP4_HSVBK	TRANS-ACTING TRANSCRIP	4.44e+01
124	84	7.5	224	1	Y091_NPVAC	HYPOTHETICAL 24.1 KD P	3.46e+01	197	82	7.4	2493	1	CYAA_USTMA	ADENYLATE CYCLASE (EC	4.44e+01
125	83	7.5	241	1	COAT_GSMV	COAT PROTEIN.	3.46e+01	198	82	7.4	4367	1	DYHC_NEUCR	DYNEIN HEAVY CHAIN, CY	4.44e+01
126	83	7.5	279	1	FR11_HUMAN	PROTO-ONCOGENE FRAT1 (	3.46e+01	199	81	7.3	268	1	CDX1_MOUSE	HOMEOBOX PROTEIN CDX-1	5.67e+01
127	83	7.5	290	1	V290_ASFLS	LIS 290 PROTEIN PRECUR	3.46e+01	200	81	7.3	289	1	EIA_ADE05	EARLY E1A 32 KD PROTEI	5.67e+01
128	83	7.5	319	1	ISL2_CHICK	INSULIN GENE ENHANCER	3.46e+01								
129	84	7.5	326	1	FR2_MOUSE	FOS-RELATED ANTIGEN 2.	2.70e+01								
130	84	7.5	347	1	VCON_ADEL12	MINOR CORE PROTEIN (PR	2.70e+01								
131	83	7.5	364	1	MYCL_HUMAN	L-MYC-1 PROTO-ONCOGENE	3.46e+01								
132	84	7.5	366	1	YJGP_ECOLI	HYPOTHETICAL 40.4 KD P	2.70e+01								
133	83	7.5	405	1	VGLM_EBV	GLYCOPROTEIN M.	3.46e+01								
134	83	7.5	427	1	TEF3_HUMAN	TRANSCRIPTIONAL ENHANC	3.46e+01								
135	84	7.5	437	1	E2F1_HUMAN	TRANSCRIPTION FACTOR E	2.70e+01								
136	83	7.5	440	1	G3PT_MOUSE	GLYCERALDEHYDE 3-PHOSP	3.46e+01								
137	83	7.5	480	1	BLAR_MACMU	BETA-1 ADRENERGIC RECE	3.46e+01								
138	83	7.5	514	1	CP24_RAT	CYTOCHROME P450-CC24 M	3.46e+01								
139	83	7.5	526	1	MSL1_YEAST	MITOCHONDRIAL GTPASE M	3.46e+01								
140	83	7.5	536	1	GAG_FSVMD	GAG POLYPROTEIN (CONTA	3.46e+01								
141	83	7.5	543	1	VP61_NPVAC	61 KD PROTEIN.	3.46e+01								
142	84	7.5	553	1	ODP2_ALCEU	DIHYDROLIPOAMIDE ACETY	2.70e+01								
143	83	7.5	556	1	HEXB_HUMAN	BETA-HEXOSAMINIDASE BE	3.46e+01								
144	83	7.5	602	1	PGH1_MOUSE	PROSTAGLANDIN G/H SYNT	3.46e+01								
145	84	7.5	615	1	MUTL_ECOLI	DNA MISMATCH REPAIR PR	2.70e+01								
146	83	7.5	628	1	V7OK_TYMYA	69 KD PROTEIN.	3.46e+01								
147	83	7.5	629	1	GIDA_HAEIN	GLUCOSE INHIBITED DIVI	3.46e+01								
148	83	7.5	629	1	GIDA_ECOLI	GLUCOSE INHIBITED DIVI	3.46e+01								
149	83	7.5	639	1	DMK_HUMAN	MYOTONIN-PROTEIN KINAS	3.46e+01								
150	83	7.5	672	1	CNAD_RAT	CAMP-DEPENDENT 3',5'-C	3.46e+01								
151	84	7.5	677	1	T2D5_HUMAN	TRANSCRIPTION INITIATI	2.70e+01								
152	84	7.5	678	1	T2D5_RAT	TRANSCRIPTION INITIATI	2.70e+01								
153	84	7.5	678	1	T2D5_MOUSE	TRANSCRIPTION INITIATI	2.70e+01								
154	83	7.5	707	1	PSF_HUMAN	PTB-ASSOCIATED SPLICIN	3.46e+01								
155	83	7.5	752	1	8511_TRYCR	SIALIDASE 85-1.1 PRECU	3.46e+01								
156	84	7.5	766	1	DOC2_MOUSE	MITOGEN-RESPONSIVE PHO	2.70e+01								
157	83	7.5	825	1	SES_RAT	SE5 ANTIGEN.	3.46e+01								
158	83	7.5	848	1	DYN3_RAT	DYNAMIN 3 (DYNAMIN, TE	3.46e+01								
159	83	7.5	915	1	YK01_YEAST	HYPOTHETICAL 102.5 KD	3.46e+01								
160	83	7.5	1083	1	T2D3_HUMAN	TRANSCRIPTION INITIATI	3.46e+01								
161	84	7.5	1091	1	DIA_DROME	DIAPHANOUS PROTEIN.	2.70e+01								
162	84	7.5	1466	1	NKCR_MOUSE	NK-TUMOR RECOGNITION P	2.70e+01								
163	83	7.5	1523	1	SON_HUMAN	SON PROTEIN (SON3).	3.46e+01								
164	83	7.5	1595	1	SOS_DROME	SON OF SEVENLESS PROTE	3.46e+01								
165	83	7.5	2241	1	TEGU_HCMV	PROBABLE LARGE TEGUMEN	3.46e+01								
166	83	7.5	2242	1	PYR1_SQUAC	CAD PROTEIN (CONTAINS:	3.46e+01								
167	83	7.5	2262	1	RRPL_P12HT	RNA POLYMERASE BETA SU	4.44e+01								
168	82	7.4	100	1	RR30_MYCTU	VERY HYPOTHETICAL 11.2	4.44e+01								
169	82	7.4	281	1	FASL_HUMAN	FAS ANTIGEN LIGAND (AP	4.44e+01								

## ALIGNMENTS

RESULT 1  
 ID IEX1\_HUMAN STANDARD; PRT; 156 AA.  
 AC P46695; Q93044; Q92691;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE RADIATION-INDUCIBLE IMMEDIATE-EARLY GENE IEX-1 (IMMEDIATE EARLY  
 DE PROTEIN GLY96) (PRG1 PROTEIN) (DIF-2 PROTEIN).  
 GN IEX1 OR PRG1 OR DIF2.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 RN EUTHERIA; PRIMATES.  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA.  
 RX MEDLINE; 96181295.  
 RA KONDRATYEV A.D., CHUNG K.-N., JUNG M.O.;  
 RL GASTROENTEROLOGY 0:0-0(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA SCHAEFER H., TRAUZOLD A., LETTAU P., KALTHOFF H., FOELSCH U.R.,  
 RA SCHMIDT W.E.;  
 RL GASTROENTEROLOGY 0:0-0(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 97394236.  
 RA PIETZSCH A., BUECHLER C., ASLANIDIS C., SCHMITZ G.;  
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 235:4-9(1997).  
 RN [4]  
 RP PRELIMINARY SEQUENCE OF 1-106 FROM N.A.  
 RC TISSUE=PLACENTA.  
 RA HILLIER L., CLARK N., DUBOQUE T., ELLISTON K., HAWKINS M., HOLMAN M.,  
 RA HULTMAN M., KUCABA T., LE M., LENNON G., WARRA M., PARSONS J.,  
 RA RIFKIN L., ROHLING T., TAN F., TREVASKIS E., WATERSTON R.,  
 RA WILLIAMSON A., WOHLDMANN P., WILSON R.;  
 RL SUBMITTED (FEB-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).  
 -1- INDUCTION: BY RADIATION, TPA, OKADAIC ACID AND TNF-ALPHA.

-1- PTM: GLYCOSYLATED.  
CC -1- SIMILARITY: STRONG, TO MOUSE ORTHOLOG.  
CC -1- CAUTION: REF.4 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO  
CC FRAMESHIFTS.  
DR EMBL: S81914; G1488385; --  
DR EMBL: X96438; E350480; --  
DR EMBL: Y14551; E333102; --  
DR EMBL: T49531; -, NOT\_ANNOTATED\_CDS.  
KW GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL-ANCHOR.  
FT DOMAIN 1 82 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 83 99 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT (POTENTIAL).  
FT DOMAIN 100 156 EXTRACELLULAR (POTENTIAL).  
FT CARBOHYD 133 133 POTENTIAL.  
FT FT CARBOHYD 54 54 A -> G (IN REF. 1).  
FT FT CONFLICT 106 106 P -> R (IN REF. 1).  
SQ SEQUENCE 156 AA; 16928 MW; 46013510 CRC32;

.. Match 100.0%; Score 1114; DB 1; Length 156;  
. Local Similarity 100.0%; Pred. No. 3,79e-194;  
tches 156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MCHSRSCPTWTIIQAAPTAPSTIPGPRSGGEIETFDLPPEAAAPACRPSASRGHRK 60  
QY 1 MCHSRSCPTWTIIQAAPTAPSTIPGPRSGGEIETFDLPPEAAAPACRPSASRGHRK 60  
Db 61 RSRLVYPVVRRQLPVVEEPNPAKRLLFLTLTVFCQILMAEGVPAPLPDPAPNAASL 120  
QY 61 RSRLVYPVVRRQLPVVEEPNPAKRLLFLTLTVFCQILMAEGVPAPLPDPAPNAASL 120  
Db 121 APTVPSPVLFPNLTPSPSYALDLSTFLOOHPPAF 156  
QY 121 APTVPSPVLFPNLTPSPSYALDLSTFLOOHPPAF 156

RESULT 2 STANDARD; PRT; 153 AA.

ID IEX1\_MOUSE  
AC P46994;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE RADIATION-INDUCIBLE IMMEDIATE-EARLY GENE IEX-1 (IMMEDIATE EARLY  
DE PROTEIN GLY96) (IMMEDIATE EARLY RESPONSE 3 PROTEIN).  
GN IEX1 OR GLY96 OR IER3.  
OS MUS MUSCULUS (MOUSE).  
OC UKAYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
CT JUTHERIA; RODENTIA.  
R. 1]  
SEQUENCE FROM N.A.  
STRAIN=BALB/C;  
MEDLINE; 93173526.  
CHARLES C.H., YOON J.K., SIMSKE J.S., LAU L.F.;  
RAL ONCOGENE 8:797-801(1993).

CC -1- FUNCTION: NOT KNOWN; EXPRESSED DURING THE GO-G1 TRANSITION OF THE  
CC CELL CYCLE.  
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).  
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE LUNG, TESTES  
CC AND THE UTERUS.  
CC -1- INDUCTION: BY SERUM GROWTH FACTORS AND TPA.  
CC -1- PTM: GLYCOSYLATED.  
CC -1- SIMILARITY: STRONG, TO HUMAN ORTHOLOG.  
DR EMBL: X67644; -, NOT\_ANNOTATED\_CDS.  
DR PIR: S33363; S33363.  
DR MGD: MG1:104814; IER3.  
KW GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL-ANCHOR.  
FT DOMAIN 1 85 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 86 102 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT (POTENTIAL).  
FT DOMAIN 103 153 EXTRACELLULAR (POTENTIAL).  
FT FT CARBOHYD 137 137 POTENTIAL.  
SQ SEQUENCE 153 AA; 16875 MW; 7A284EC2 CRC32;

Query Match 67.5%; Score 752; DB 1; Length 153;

Query Match 10.1% Score 112; DB 1; Length 228;  
 Best Local Similarity 29.3%; Pred. No. 1.01e-02;  
 Matches 27; Conservative 21; Mismatches 41; Indels 3; Gaps 3;

DB 123 PPAPVPPVGVVHYHSIEVARPPARISSPPR-RRRRRRSPRPPATAAYRSAAEV 180  
 QY 21 PSTIPGRGSGEIFTFDLPFAAPAGRPASGRHRRSRVLYPRVRRQLPVEEP 80  
 DB 181 VERRRRVAQTPVVRVHPSIQVEPAVHPPLAP 212  
 QY 81 NPAKRLFLLLTIV-FCQILMAEGVPAPLPP 111

RESULT 5  
 ID Y025-CABEL STANDARD; PRT: 1799 AA.

AC P34675;  
 01-FEB-1994 (REL. 28, CREATED)  
 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 \*PHOTHEICAL 202.6 KD PROTEIN ZK688.5 IN CHROMOSOME III.  
 C: 688.5.

OS AENORHABDITIS ELEGANS.  
 OC EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPESEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FRASER A.,  
 RA FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,  
 RA JOHNSON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N.,  
 RA LARSON J., LIGHTNING J., LLOYD C., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHONKKEEN R.,  
 RA SIMS M., SKALDON N., SMITH A., SMITH M., SONNHAMMER E., STADEN R.,  
 RA SULTON J., THIERY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K.,  
 RA WATSON R., WATSON A., WEINSTOCK L., WILKINSON-SPROAT J.,  
 RA WOHLDMAN P.;  
 RL NATURE 368:32-38(1994).  
 DR EMBL; L16621; G289783; -.  
 DR PIR; S44920; S44920.  
 DR WORMPEP; ZK688.5; CE00463.  
 DR PROSITE; P550053; UBIQUITIN\_2; UNKNOWN\_1.  
 KW HYPOTHETICAL PROTEIN.  
 FT DOMAIN 21 96 UBIQUITIN-LIKE.  
 SEQUENCE 1799 AA; 202641 MW; 35F7DDB2 CRC32;

Query Match 9.8% Score 109; DB 1; Length 1799;  
 Local Similarity 23.7%; Pred. No. 2.51e-02;  
 Matches 31; Conservative 30; Mismatches 65; Indels 5; Gaps 5;

DB 844 FLENGRIPSTSSAPSTSENPPGPFNS-EDAADIRAGRLPLGTGRNRRRTVRETVHAAA 902  
 QY 13 ILQAPTAPSTIPGRGSGEIFTFDLPFAAPAGR-PSASGRHRRSRVLYPRV 71  
 DB 903 ARAESPNHSLITATHTHFAAGF-PLMASSNVSTSGAGPCWIRQVQVSTPTTRGL 961  
 QY 72 RR-OLPVEENPAKRLFLLLTIVFCQILMAEGVPAPLP-PEDAPNAASLAPTVPV 129  
 DB 962 FEPDLGSSDQ 972  
 QY 130 EPNLISEPD 140

RESULT 6  
 ID FOR4-MOUSE STANDARD; PRT: 1206 AA.

AC Q05859;  
 01-JUN-1994 (REL. 29, CREATED)  
 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
 DE FORMIN 4 (LIMB DEFORMITY PROTEIN).  
 GN LD.

MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-EMBRYO;  
 RX MEDLINE; 92112033.  
 RA GRUBBY-JACKSON L., KUO A., LEDER P.;  
 RL GENES DEV. 6:29-37(1992).  
 CC FUNCTION: IS IMPORTANT IN THE MORPHOGENESIS OF LIMB AND MAY HAVE A  
 CC FUNCTION IN DIFFERENTIATED CELLS OR BE INVOLVED IN MAINTAINING  
 CC SPECIFIC DIFFERENTIATED STATES.  
 CC TISSUE SPECIFICITY: IT IS FOUND THROUGHOUT THE EMBRYO BUT  
 CC HAS A FUNCTIONAL ROLE ONLY IN THE KIDNEY AND LIMB.  
 CC DEVELOPMENTAL STAGE: THIS IS THE ISOFORM FOUND IN THE APICAL  
 CC ECTODERMAL RIDGE AND THE MESENCHYMAL COMPARTMENT OF THE DEVELOPING  
 CC LIMB BUD.  
 CC ALTERNATIVE PRODUCTS: MANY DIFFERENT FORMS OF THIS PROTEIN ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING OF THE LD GENE. A VARIATION IN  
 CC SPLICING IS SEEN AMONG DIFFERENT TISSUES AND DIFFERENT SIZE  
 CC TRANSCRIPTS EXIST WITHIN ANY ONE TISSUE. THIS IS ISOFORM 4 AND IS  
 CC DIFFERENT IN ITS N-TERMINAL TO THE OTHER MOUSE ISOFORMS SO FAR  
 CC DETERMINED.  
 DR EMBL; X52379; G51553; -.  
 DR PIR; S24407; S24407.  
 DR HSP; P19999; ICLUG.  
 KW NUCLEAR PROTEIN; DEVELOPMENTAL PROTEIN; ALTERNATIVE SPLICING.  
 FT DOMAIN 635 638 POLY-SER.  
 FT DOMAIN 644 744 PRO-RICH.  
 FT DOMAIN 751 755 POLY-SER.  
 SQ SEQUENCE 1206 AA; 133464 MW; 6D70C261 CRC32;  
 Query Match 9.2% Score 103; DB 1; Length 1206;  
 Best Local Similarity 39.1%; Pred. No. 1.50e-01;  
 Matches 18; Conservative 15; Mismatches 9; Indels 4; Gaps 4;

DB 716 VLALNSGCGPP-PVPPPPPPGLAPP-PPGL-SFGLSSSSSQYP 758  
 QY 98 IL-MAEGVPAPLPEDAPNAASLAPTVPVLEPFLNLTSEPSDYA 142

RESULT 7

ID FORM-MOUSE STANDARD; PRT: 1468 AA.

AC Q05860;  
 DT 01-JUN-1994 (REL. 29, CREATED)  
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE FORMIN (LIMB DEFORMITY PROTEIN).  
 GN FMN OR LD.

OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE-KIDNEY, AND TESTIS;  
 RX MEDLINE; 90363291.  
 RA WOCHIK R.P., MAAS R.L., ZELLER R., VOGT T.F., LEDER P.;

RL NATURE 346:850-853(1990).  
 CC FUNCTION: IS IMPORTANT FOR THE MORPHOGENESIS OF LIMB AND KIDNEY  
 CC AND MAY HAVE A FUNCTION IN DIFFERENTIATED CELLS OR MAY BE  
 CC INVOLVED IN MAINTAINING SPECIFIC DIFFERENTIATED STATES.  
 CC TISSUE SPECIFICITY: IT IS PRESENT IN THE ADULT KIDNEY, TESTIS,  
 CC LIMB, OVARY, BRAIN, SMALL INTESTINE, SALIVARY GLAND AND HARDERIAN  
 CC GLAND. IT IS PRESENT THROUGHOUT THE EMBRYO.  
 CC DEVELOPMENTAL STAGE: IN THE DEVELOPING LIMB BUD, THE PROTEIN  
 CC IS EXPRESSED IN THE APICAL ECTODERMAL RIDGE AND THE MESENCHYMAL  
 CC COMPARTMENT, PREDOMINANTLY IN THE POSTERIOR REGION. DURING  
 CC KIDNEY MORPHOGENESIS, EXPRESSION IS INITIALLY RESTRICTED TO  
 CC THE EPITHELIAL COMPARTMENT OF THE PRONEPHROS AND MESONEPHROS.  
 CC ALTERNATIVE PRODUCTS: MANY DIFFERENT FORMS OF THIS PROTEIN  
 CC ARE PRODUCED BY ALTERNATIVE SPLICING OF THE LD GENE. A  
 CC VARIATION IN SPLICING IS SEEN AMONG DIFFERENT TISSUES AND

\*\*\*\*\*

W A R E S R E H

(TW)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm  
b-- On: Sun Jan 10 02:36:53 1999; Maspar time 2001.26 Seconds  
1352.279 Million cell updates/sec  
lar output not generated.

Title: >US-08-799-910-9  
Description: (1-1228) from US08799910.seq  
Perfect Score: 1228  
N.A. Sequence: 1 ATGTGTCACCTCGCAGCTG.....AAAAA.....AACTCGAG 1228  
Comp: TACACAGTGAGCGCTCGAC.....TTTTTTTTTTTGAGCTC

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 567134 seqs, 1101898692 bases x 2

Post-processing: Minimum Match 0%  
Listing first 200 summaries

Database: emb155  
1:em\_ba 2:em\_fun 3:em\_htg 4:em\_hum1 5:em\_hum2 6:em\_in  
7:em\_om 8:em\_or 9:em\_ov 10:em\_pat 11:em\_ph 12:em\_pl  
13:em\_ro 14:em\_vi  
genbank107  
15:gb\_bal 16:gb\_ba2 17:gb\_htg 18:gb\_in 19:gb\_om 20:gb\_ov  
21:gb\_pat 22:gb\_ph 23:gb\_pl 24:gb\_pl2 25:gb\_pr1  
26:gb\_pr2 27:gb\_pr3 28:gb\_ro 29:gb\_st 30:gb\_sts 31:gb\_sy  
32:gb\_un 33:gb\_vi

clistics: Mean 11.515; Variance 7.482; scale 1.539

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1205	98.1	1230	26	HSDIF2 Homo sapiens mRNA for	0.00e+00
2	1171	95.4	1223	25	TEX-1-radiation-Induc1	0.00e+00
3	976	79.5	1309	27	AF039067 Homo sapiens anti-deat	0.00e+00
4	917	74.7	1864	26	HSPRG1 H.sapiens PRG1 gene.	0.00e+00
5	471	38.4	477	27	AF083421 Homo sapiens radiation	7.53e-276
6	313	25.5	343	26	HSA227914 Homo sapiens partial m	3.31e-173
7	118	9.6	1938	28	MMGLY96 M.musculus gly96 mRNA.	6.27e-50
8	114	9.3	1758	28	RNRPG1 R.norvegicus PRG1 gene	1.67e-47
9	79	6.4	7218	21	166494 Sequence 14 from paten	9.39e-27
10	67	5.5	7218	21	166494 Sequence 14 from paten	6.44e-20
11	41	3.3	215	21	128278 Sequence 5 from patent	4.73e-06
12	39	3.2	1368	26	HS16736 Homo sapiens dif-2 gen	4.53e-05
13	38	3.1	74371	26	AC005369 Homo sapiens chromosom	1.39e-04

14	10772	18	AF012089	Drosophila melanogaste	4.16e-04
15	10772	18	AF012089	Drosophila melanogaste	3.65e-03
16	74371	26	AC005369	Homo sapiens chromosom	1.24e-03
17	192	30	H0451251	Human Chromosome 4 (cl	1.06e-02
18	408	30	G34020	Sequence 5 from patent	3.04e-02
19	215	21	128278	Human U1 small nuclear	2.39e-01
20	1662	25	H0MRNP70A	Mus musculus glycosylp	2.39e-01
21	3436	28	AF050666	Gallus gallus fibrobla	1.76e-00
22	800	20	GGU41467	S.scrofa mRNA for L-36	1.76e-00
23	1125	19	SSL36LBP	Bovine inorganic pyrop	1.76e-00
24	1266	19	BOVIOFPP	Dirofilaria immitis 70	1.76e-00
25	1294	18	DIRHSP70	Mus musculus mRNA for	1.76e-00
26	1479	28	NN26SPROT	Xenopus laevis nuclear	1.76e-00
27	1811	20	XU595669	Dictyostelium discoide	1.76e-00
28	1840	18	DDU73686	H.sapiens mig-5 gene.	1.76e-00
29	1941	25	HSNIGGE	Rana catesbeiana allan	1.76e-00
30	2116	20	RCALN	Danio rerio stem cell	1.76e-00
31	2873	19	AF045432	S.scrofa mRNA for inte	1.76e-00
32	3042	19	SS11545	Zea mays LON1 protease	1.76e-00
33	5433	28	AF090113	Cyl-1-cyclin D1 (mice,	1.76e-00
34	3217	24	ZMU85494	Xenopus laevis myb-rel	6.55e-01
35	3737	28	S78355	Homo sapiens Cdc14B1 p	1.76e-00
36	3789	20	XELMYBRP1	Rattus norvegicus AMPA	1.76e-00
37	4624	26	AF064104	*** SEQUENCING IN PROG	1.76e-00
38	5433	28	AF090113	human STS SHGC-15668	4.65e-00
39	224645	17	AC004153	SHGC-57073 Human Homo	4.65e-00
40	232	30	G15123	Drosophila melanogaste	4.65e-00
41	350	30	G37172	Drosophila melanogaste	4.65e-00
42	571	18	SCU41441	Homo sapiens full leng	4.65e-00
43	648	26	HUMYH1B08	laminin receptor homol	4.65e-00
44	739	25	S35960	Homo sapiens clone 246	4.65e-00
45	1332	26	AF070648	casein kinase 2 alpha	4.65e-00
46	1404	20	S76875	Drosophila melanogaste	4.65e-00
47	1435	18	DMU75652	Human clone 23960 mRNA	4.65e-00
48	1659	26	HSU79276	Homo sapiens clone 245	4.65e-00
49	1841	26	AF070588	M.musculus seb4 mRNA.	4.65e-00
50	1908	28	MNRNASEB4	Mouse mRNA for prothro	4.65e-00
51	2031	28	MMPTTHROM	CDNA encoding alpha 2,	4.65e-00
52	2179	21	E08204	Rattus norvegicus UDP-	4.65e-00
53	2301	28	RNU75903	Dictyostelium discoid	4.65e-00
54	2624	21	AF020409	Sequence 1 from patent	4.65e-00
55	2628	21	I66342	Homo sapiens tyrosine	4.65e-00
56	2646	26	AF023158	Homo sapiens transcrip	4.65e-00
57	3946	26	AF048693	Dictyostelium discoid	4.65e-00
58	5061	18	DDU25144	Human trans-Golgi p230	4.65e-00
59	7695	27	HSU41740	Caenorhabditis elegans	4.65e-00
60	38890	18	CELC14C11	S.pombe chromosome II	4.65e-00
61	42934	23	SPBC347	Arabidopsis thaliana g	4.65e-00
62	58527	26	AB007649	Homo sapiens DNA sequ	4.65e-00
63	118226	26	HS46618	Caenorhabditis elegans	4.65e-00
64	292720	17	CEY47H10	human STS WI-15413.	1.21e+01
65	246	30	G21629	human STS WI-13630.	1.21e+01
66	474	30	G24648	Mus musculus nuclear o	1.21e+01
67	479	28	NMRORC5	Mus musculus long int	1.21e+01
68	500	28	MMSLIB44	B.verrucosa Bet v lg m	1.21e+01
69	701	23	BVBETVLG	B.verrucosa Bet v lg m	1.21e+01
70	714	23	BVBETVLB	A.thaliana mRNA for Cu	1.21e+01
71	788	23	ATCZ5UD1	Danio rerio mRNA for d	1.21e+01
72	802	20	AB005204	Mus musculus probasin	1.21e+01
73	884	28	PRDB428	Rattus norvegicus pren	1.21e+01
74	884	28	PRDB428	P.reticulata Mhc gene,	1.21e+01
75	907	20	AF012463	Pleuronectes americanu	1.21e+01
76	945	23	AF050102	Oryza sativa glutathio	1.21e+01
77	977	20	XLXVENT1	X.laevis mRNA for Xven	1.21e+01
78	979	20	S78789S4	CGATA-3 [chickens, liv	1.21e+01
79	1216	18	SCDCD2A	C.elegans mRNA for p34	1.21e+01
80	1299	21	I76370	Sequence 1 from patent	1.21e+01
81	1362	23	PSY14273	Pisum sativum mRNA for	1.21e+01
82	1444	18	HELWAMIDE	H.echinata mRNA for LW	1.21e+01
83	1452	25	HSU63108	Human eukaryotic initi	1.21e+01
84	1588	27	HUMORF1	Homo sapiens (clone S2	1.21e+01
85	1727	26	AF054999	Homo sapiens clone 247	1.21e+01
86	2240	18	LCU56636	Lucilia cuprina alpha	1.21e+01



87	27	2.2	2295 18	AF020901	Microclona prolifera a	1.21e+01	160	26	2.1	3254 20	XXXSRC	X.xiphidium Xsrc mRNA.	3.06e+01
88	27	2.2	2305 26	HSU81787	Human Ant10B mRNA, com	1.21e+01	161	26	2.1	3268 18	PFRESAG1	P.falciapaur FC27 RESA	3.06e+01
89	27	2.2	2498 28	MAU43333	Mesocricetus auratus 5	1.21e+01	162	26	2.1	3281 20	GGU2964	Gallus gallus beta-cat	3.06e+01
90	27	2.2	2818 23	BLYL0XA	Barley lipoxigenase 1	1.21e+01	163	26	2.1	3341 25	HSU51224	Human U2AFBFL gene, co	3.06e+01
91	27	2.2	2842 21	A26375	RING4 coding sequence.	1.21e+01	164	26	2.1	3767 27	AF084530	Homo sapiens cyclin-D	3.06e+01
92	27	2.2	3016 18	DDU31631	Dictyostelium discoidie	1.21e+01	165	26	2.1	3778 28	AF006688	Mus musculus peroxisom	3.06e+01
93	27	2.2	3819 21	DDU31631	Sequence 3 from patent	1.21e+01	166	26	2.1	3936 25	HSADA	Human mRNA for erythro	3.06e+01
94	27	2.2	4336 26	AF027299	Homo sapiens protein 4	1.21e+01	167	26	2.1	3950 31	EVU81273	Expression vector pXen	3.06e+01
95	27	2.2	6265 18	PFSC03080	Plasmodium falciapaur	1.21e+01	168	26	2.1	3966 31	EVU81274	Expression vector pXen	3.06e+01
96	27	2.2	8595 18	DDU14576	Dictyostelium discoidie	1.21e+01	169	26	2.1	3970 31	EVU81275	Expression vector pXen	3.06e+01
97	27	2.2	10168 18	PFSC03091	Plasmodium falciapaur	1.21e+01	170	26	2.1	3973 21	I32039	Sequence 21 from paten	3.06e+01
98	27	2.2	14867 18	PFSC03072	Plasmodium falciapaur	1.21e+01	171	26	2.1	4007 26	AF000670	Homo sapiens elf-1 rel	3.06e+01
99	27	2.2	24741 18	CELC44C3	Caenorhabditis elegans	1.21e+01	172	26	2.1	9113 28	AB011678	Mus musculus mRNA for	3.06e+01
100	27	2.2	65691 18	PMAL3P1	Plasmodium falciapaur	1.21e+01	173	26	2.1	10484 33	FSU85043	Feline syncytial virus	3.06e+01
101	27	2.2	92879 17	AC004450	*** SEQUENCING IN PROG	1.21e+01	174	26	2.1	10227 18	CEC01C7	Caenorhabditis elegans	3.06e+01
102	27	2.2	108245 26	AC005166	Homo sapiens PAC clone	1.21e+01	175	26	2.1	25400 18	CEFS9A2	Caenorhabditis elegans	3.06e+01
103	27	2.2	122940 17	AC004709	*** SEQUENCING IN PROG	1.21e+01	176	26	2.1	40387 23	SPAC23H4	S.pombe Chromosome I c	3.06e+01
104	27	2.2	147394 27	HS353H6	Human DNA sequence fro	1.21e+01	177	26	2.1	42759 24	AB004535	Schizosaccharomyces po	3.06e+01
105	27	2.2	155622 17	AC004839	*** SEQUENCING IN PROG	1.21e+01	178	26	2.1	69964 27	HS496C20	Human DNA sequence fro	3.06e+01
106	27	2.2	189279 27	AC005028	Homo sapiens BAC clone	1.21e+01	179	26	2.1	107135 26	HS145B12	Human DNA sequence fro	3.06e+01
107	27	2.2	199878 17	AC005027	*** SEQUENCING IN PROG	1.21e+01	180	26	2.1	113950 24	ATAC004697	Arabidopsis thaliana c	3.06e+01
108	27	2.2	224645 17	AC004153	*** SEQUENCING IN PROG	1.21e+01	181	26	2.1	115355 17	CEY55D9	Caenorhabditis elegans	3.06e+01
109	26	2.1	160 23	DIATIE	D.discoidium actin A-3	3.06e+01	182	26	2.1	122000 17	AC004050	*** SEQUENCING IN PROG	3.06e+01
110	26	2.1	308 21	I80066	Sequence 40 from paten	3.06e+01	183	26	2.1	128915 26	AC002416	Human Chromosome X, co	3.06e+01
111	26	2.1	451 18	MSLLRN	S.latinus mitochondr	3.06e+01	184	26	2.1	129169 17	AC005506	*** SEQUENCING IN PROG	3.06e+01
112	26	2.1	463 30	G13251	Human STS WI-12507.	3.06e+01	185	26	2.1	131070 27	AC005295	Homo sapiens Xp22 bins	3.06e+01
113	26	2.1	465 18	AF069172	Euphidius setiger NAD	3.06e+01	186	26	2.1	139097 17	AC004591	*** SEQUENCING IN PROG	3.06e+01
114	26	2.1	541 23	MD080271	Malus domestica prolin	3.06e+01	187	26	2.1	136601 26	HS21266	Homo sapiens DNA seque	3.06e+01
115	26	2.1	549 23	HYSC12	H.vulgare mRNA for sub	3.06e+01	188	26	2.1	160262 26	AC005242	Homo sapiens Chromosom	3.06e+01
116	26	2.1	614 28	MUSUNKK	Mouse (clone HORT11) m	3.06e+01	189	26	2.1	163847 17	AC004821	*** SEQUENCING IN PROG	3.06e+01
117	26	2.1	707 24	LPJ5047	Lycopersicon peruvianu	3.06e+01	190	26	2.1	172533 26	AC002381	Human BAC clone RG020D	3.06e+01
118	26	2.1	1017 21	A62439	Sequence 1 from Patent	3.06e+01	191	26	2.1	193374 17	AC005139	*** SEQUENCING IN PROG	3.06e+01
119	26	2.1	1017 21	A62449	Sequence 1 from Patent	3.06e+01	192	26	2.1	203583 17	AC004940	*** SEQUENCING IN PROG	3.06e+01
120	26	2.1	1132 28	RNRNASC33	R.norvegicus mRNA for	3.06e+01	193	26	2.1	216021 26	HUAC004787	Homo sapiens Chromosom	3.06e+01
121	26	2.1	1170 23	OSU65956	Oryza sativa GF14-b pr	3.06e+01	194	26	2.1	216021 26	HUAC004787	Homo sapiens Chromosom	3.06e+01
122	26	2.1	1185 23	ATCLAPSK	A.thaliana (L.Heynh.)	3.06e+01	195	26	2.1	239737 17	HS353E16	Human DNA sequence ***	3.06e+01
123	26	2.1	1190 18	ST5262	ncc-1-p34cdc2 kinase (	3.06e+01	196	26	2.1	244649 17	AC005308	*** SEQUENCING IN PROG	3.06e+01
124	26	2.1	1284 20	XEHAND	X.laevis mRNA for eHAN	3.06e+01	197	26	2.1	244649 17	AC005308	*** SEQUENCING IN PROG	3.06e+01
125	26	2.1	1326 28	S0547	procathepsin E [guinea	3.06e+01	198	26	2.1	245439 28	MWHC135G15	Mus musculus major his	3.06e+01
126	26	2.1	1326 28	GFPROCAT	Cavia porcellus procat	3.06e+01	199	26	2.1	251364 28	MWAE000564	Mus musculus TCR beta	3.06e+01
127	26	2.1	1349 21	I41429	Sequence 20 from paten	3.06e+01	200	26	2.1	259894 26	HUAC002302	Homo sapiens Chromosom	3.06e+01
128	26	2.1	1364 19	AF060221	Sus scrofa deoxyribonu	3.06e+01							
129	26	2.1	1364 18	MOTPA4	Manduca sexta hemolin	3.06e+01							
130	26	2.1	1373 23	LETHM1	L.esculentum mRNA for	3.06e+01							
131	26	2.1	1390 21	A43707	Sequence 4 from Patent	3.06e+01							
132	26	2.1	1393 26	D55425	Human mRNA for transac	3.06e+01							
133	26	2.1	1404 21	A43736	Sequence 2 from Patent	3.06e+01							
134	26	2.1	1472 18	MDP3	M.domestica mRNA for y	3.06e+01							
135	26	2.1	1485 21	I75251	Sequence 23 from paten	3.06e+01							
136	26	2.1	1485 21	I75226	Sequence 23 from paten	3.06e+01							
137	26	2.1	1490 21	A43708	Sequence 3 from Patent	3.06e+01							
138	26	2.1	1498 28	RNU68844	Rattus norvegicus cycl	3.06e+01							
139	26	2.1	1778 28	RNU48219	Rattus norvegicus cyto	3.06e+01							
140	26	2.1	1833 26	HSU79290	Human clone 23908 mRNA	3.06e+01							
141	26	2.1	1853 26	AF070609	Homo sapiens clone 246	3.06e+01							
142	26	2.1	1940 23	AF070609	Lycopersicon esculentu	3.06e+01							
143	26	2.1	2048 28	MUSNPRCC	Mouse mRNA for nuclear	3.06e+01							
144	26	2.1	2056 27	HSU50527	Human BRCA2 region, mR	3.06e+01							
145	26	2.1	2152 28	CCPH20	Cavia cobyaya mRNA for	3.06e+01							
146	26	2.1	2152 21	I89388	Sequence 1 from patent	3.06e+01							
147	26	2.1	2152 25	HSEPI1	H.sapiens mRNA for epi	3.06e+01							
148	26	2.1	2173 21	I32035	Sequence 6 from patent	3.06e+01							
149	26	2.1	2178 26	AF055008	Homo sapiens clone 247	3.06e+01							
150	26	2.1	2198 28	AB007812	Mus musculus mGLT-1B m	3.06e+01							
151	26	2.1	2370 27	HSU48596	Human mariner-like ele	3.06e+01							
152	26	2.1	2394 27	HSU87166	Homo sapiens spectrin	3.06e+01							
153	26	2.1	2692 21	I32038	Sequence 14 from paten	3.06e+01							
154	26	2.1	2722 20	CHRWYOSINA	Chicken fast myosin he	3.06e+01							
155	26	2.1	2812 20	XLJ92030	Xenopus laevis TGF-bet	3.06e+01							
156	26	2.1	2892 23	DDISGSPA	Dictyostelium discoidie	3.06e+01							
157	26	2.1	3015 26	MFPAZMRNA	Macaca fascicularis mR	3.06e+01							
158	26	2.1	3073 20	GGU07775	Gallus gallus prepro-7	3.06e+01							
159	26	2.1	3163 25	HSDYNEIN	H.sapiens mRNA dynein-	3.06e+01							

ALIGNMENTS

RESULT	1	HSDF2	1230 bp	RNA	PRI	12-AUG-1997
LOCUS		Homo sapiens mRNA for DIF-2 protein.				
DEFINITION						
ACCESSION		Y14551				
NID		g2660550				
KEYWORDS		dif-2 gene.				
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
AUTHORS		Primates; Catarrhini; Hominiidae; Homo.				
TITLE		1 (bases 1 to 1230)				
		Pietzsch, A., Buchler, C., Aslanidis, C. and Schmitz, G.				
		Identification and characterization of a novel monocyte/macrophage				
		differentiation-dependent gene that is responsive to				
		lipopolysaccharide, ceramide, and lysophosphatidylcholine				
		Biochem. Biophys. Res. Commun. 235 (1), 4-9 (1997)				
JOURNAL		97339426				
MEDLINE		2 (bases 1 to 1230)				
REFERENCE						
AUTHORS		Pietzsch, A.				
TITLE		Direct Submission				
JOURNAL		Submitted (11-AUG-1997) A. Pietzsch, Institute for Clinical				
		Chemistry and Laboratory Medicine, University of Regensburg,				
		Franz-Josef-Strauss-Allee 11, 93053 Regensburg, FRG				
REFERENCE		3 (bases 1 to 1230)				
AUTHORS		Kondratyev, A.D., Chung, K.N. and Jung, M.O.				
TITLE		Identification and characterization of a radiation-inducible				
JOURNAL		glycosylated human early-response gene				
		Cancer res. 56 (7), 1498-1502 (1996)				



MEDLINE	96181295	Location/Qualifiers
FEATURES		
Source		1..1230 /organism="Homo sapiens" /db_xref="taxon:9606" /cell_type="leukocytes"
5'UTR		<1..11 /gene="dif-2"
gene		1..1230 /gene="dif-2"
CDS		12..482 /gene="dif-2"
		/codon_start=1 /product="Dif-2 protein" /db_xref="PID:e333102" /db_xref="PID:g2660551"
		/translation="MCHSRSCHPTWILQAPTPAPSTIPGPRRGSGPEITFDPLPPEP AAAPAGPSASGHKRSRVLYPVVRVROLVPEPNPAKRLLFILLTLTVFCQLLMAE EGVPAPLPDPEDAPNAASLAPTVPSPVLEFPNLTSEPSDYALDLSTFLQHHPAF"
conflict		172 /gene="dif-2" /citation={3} /replace="g"
conflict		328 /gene="dif-2" /citation={3} /replace="g"
3'UTR		483..>1230 /gene="dif-2"
conflict		673..675 /gene="dif-2" /citation={3} /replace="ag"
conflict		687..689 /gene="dif-2" /citation={3}
conflict		783..785 /gene="dif-2" /citation={3} /replace="tg"
conflict		1020..1021 /gene="dif-2" /citation={3}
conflict		1120 /gene="dif-2" /citation={3} /replace="att"
conflict		1133 /gene="dif-2" /citation={3} /replace="g"
BASE COUNT	259 a 367 c 342 g 262 t	
ORIGIN		
		Query Match 98.1%; Score 1205; DB 26; Length 1230; Best Local Similarity 99.4%; Pred.No. 0.00e+00; Matches 1212; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Dd	12 ATGTGTCACCTCGCAGCTGCCACCGACCATGACCATCTGCAGGGCCGACCCCGGCC 71	
Qy	1 ATGTGTCACCTCGCAGCTGCCACCGACCATGACCATCTGCAGGGCCGACCCCGGCC 60	
Dd	72 CCCTCCACCATCCGGGACCCCGGGGCTCCGGTCTTGAGATCTTCACCTTCGACCCCT 131	
Qy	61 CCCTCCACCATCCGGGACCCCGGGGCTCCGGTCTTGAGATCTTCACCTTCGACCCCT 120	
Dd	132 CTCCCGGAGCCGCGAGCGGCCCTGCGGGCGGCCCGCACGCCCTCTCGCGGGCACCCGAAG 191	
Qy	121 CTCCCGGAGCCGCGAGCGGCCCTGCGGGCGGCCCGCACGCCCTCTCGCGGGCACCCGAAG 180	
Dd	192 CGCAGCCGCAAGGGTTCTACTCCCTCGAGTGGTCCGGGGCCGACGTGCCAGTCGAGAACCG 251	

LOCUS S81914 1223 bp mRNA PRI 12-AUG-1996  
 DEFINITION IEX-1-radiation-inducible immediate-early gene [human, placenta, mRNA Partial, 1223 nt].  
 ACCESSION S81914  
 NID 9148384  
 KEYWORDS human placenta.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryotes; Chordata; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1223)  
 AUTHORS Kondratyev, A.D., Chung, K.N. and Jung, M.O.  
 TITLE Identification and characterization of a radiation-inducible glycosylated human early-response gene  
 JOURNAL Cancer Res. 56 (7), 1498-1502 (1996)  
 MEDLINE 96181295  
 REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gisbq 17109] from the original journal article.  
 :URES This sequence comes from Fig. 2A.  
 :source Location/Qualifiers  
 1..1223  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 19..489  
 /gene="IEX-1"  
 19..489  
 /gene="IEX-1"  
 /note="radiation-inducible immediate-early gene"  
 /note="radiation-inducible immediate-early gene; This sequence comes from Fig. 2A"  
 19..489  
 /codon\_start=1  
 /db\_xref="PIR:g1488385"  
 /translation="MCHSRCHPTMTILOAPTAPSTIPGRSGPEIFTDPLEP  
 AAPAGPSSRGHRKSRVLYPRVVRQLPVEENPAKRLFLLLTIVFCQILAAE  
 EGVRAFPFDDAANAALATPVSPLFENLSEPSDYALDLSLTLQHPAAF"  
 BASE COUNT 249 a 368 c 342 g 264 t  
 ORIGIN  
 Query Match 95.4%; Score 1171; DB 25; Length 1223;  
 Best Local Similarity 99.4%; Pred. No. 0.00e+00;  
 Matches 1198; Conservative 0; Mismatches 3; Indels 4; Gaps 4;  
 Db 19 ATGTGTACTCTCGAGCTGCGACCGACCATGACCATCTCGAGCGCCCGACCCCGGCC 78  
 Qy 1 ATGTGTACTCTCGAGCTGCGACCGACCATGACCATCTCGAGCGCCCGACCCCGGCC 60  
 P 79 CCCTCCACATCCCGGACCGCGGGGGCTCCGGTCTGAGATCTTCACCTTCGACCT 138  
 C 61 CCCTCCACATCCCGGACCGCGGGGGCTCCGGTCTGAGATCTTCACCTTCGACCT 120  
 Db 139 CTCGGGAGCGGACCGCGCCCTCGCGGGCGCCCGACCGCTCTCGGGGCGACCGAAAG 198  
 Qy 121 CTCGGGAGCGGACCGCGCCCTCGCGGGCGCCCGACCGCTCTCGGGGCGACCGAAAG 180  
 Db 199 CGACCGCAGGGTTCTTACCTCTGAGTGTGCGGCGCCAGCTGCCAGTCCGAGGAACCG 258  
 Qy 181 CGACCGCAGGGTTCTTACCTCTGAGTGTGCGGCGCCAGCTGCCAGTCCGAGGAACCG 240  
 Db 259 AACCCAGCAAAAGCTTCTTTCTGCTGCTCACCATCGTCTTCGCCAGATCTGTATG 318  
 Qy 241 AACCCAGCAAAAGCTTCTTTCTGCTGCTCACCATCGTCTTCGCCAGATCTGTATG 300  
 Db 319 GCTGAAGAGGGTGTGCGGCGCCCTCGCTCCAGAGGACGCCCTTAACCGCGCATCCCTG 378  
 Qy 301 GCTGAAGAGGGTGTGCGGCGCCCTCGCTCCAGAGGACGCCCTTAACCGCGCATCCCTG 360  
 Db 379 GCGCCACCCCTGTGTCCCGCGTCTCGAGCCCTTTAATCTGACTTCGGAGCCCTCGGAC 438  
 Qy 361 GCGCCACCCCTGTGTCCCGCGTCTCGAGCCCTTTAATCTGACTTCGGAGCCCTCGGAC 420  
 Db 439 TAGCTCTGGAGCTCAGACACTTTCCTCAGCAACACCGCGGCGCTTCTACTGTGACTC 498  
 Qy 421 TAGCTCTGGAGCTCAGACACTTTCCTCAGCAACACCGCGGCGCTTCTACTGTGACTC 480

Db 499 CCGGCACTCCCAAAAGAAATCGAAAAACCAAAAGAAACACGAGCGTACCTGTGGC 558  
 Qy 481 CCGGCACTCCCAAAAGAAATCGAAAAACCAAAAGAAACACGAGCGTACCTGTGGC 540  
 Db 559 CGAGAGCGTATCCCAACTCCGAGCTTCCGAGCAACTTGAATCAGAACACACACACGGA 618  
 Qy 541 CGAGAGCGTATCCCAACTCCGAGCTTCCGAGCAACTTGAATCAGAACACACACACGGA 600  
 Db 619 GAGCCACCGGGTCTTGGAGCGGACCGAGGCGACAGAGCGGAGGCGCATAGAGACC 678  
 Qy 601 GAGCCACCGGGTCTTGGAGCGGACCGAGGCGACAGAGCGGAGGCGCATAGAGACC 660  
 Db 679 GAG-CACAGCGCCAGCTGGG-CTAGGCGCGGTGGGAAGAGAGCGTCTTAATTTATTTCT 736  
 Qy 661 GAGGACAGCGCCAGCTGGGCTAGGCGTGGGAAGAGAGCGTCTTAATTTATTTCT 720  
 Db 737 TATTGCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTATTTCT 795  
 Qy 721 TATTGCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTATTTCT 780  
 Db 796 GTACGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTATTTCT 855  
 Qy 781 GTACGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTATTTCT 840  
 Db 856 AGCTCTCTTGGTATTTATTGAGCTTTTGGGACTTTGGGAGCTTTGGGAGCAGGACACCTGGAATGC 915  
 Qy 841 AGCTCTCTTGGTATTTATTGAGCTTTTGGGACTTTGGGAGCAGGACACCTGGAATGC 900  
 Db 916 GGCAAGTAGGAGAGAAATGGGAGGACTCGGGTGGGAGGAGGAGCTCCGGCTGGGATG 975  
 Qy 901 GGCAAGTAGGAGAGAAATGGGAGGACTCGGGTGGGAGGAGGAGCTCCGGCTGGGATG 960  
 Db 976 AACTCTGGTGGTGGTGGTAAAGTTAGGAGTGCATCTCCAGCATCTCAACTCC 1035  
 Qy 961 AACTCTGGTGGTGGTGGTAAAGTTAGGAGTGCATCTCCAGCATCTCAACTCC 1019  
 Db 1036 GTCTGTCTACTGTGAGACTTCGGGAGCATTAGGAATGAGATCCGTGAGATCTCTTC 1095  
 Qy 1020 GTCTGTCTACTGTGAGACTTCGGGAGCATTAGGAATGAGATCCGTGAGATCTCTTC 1079  
 Db 1096 ATCTTCTGAAGTCGCTTTAGGTTGGCTCGAGGTTAGAGGTTGGGGTTGGTGGCTG 1155  
 Qy 1080 ATCTTCTGAAGTCGCTTTAGGTTGGCTCGAGGTTAGAGGTTGGGGTTGGTGGCTG 1139  
 Db 1156 TCAGGAGCGACTGTGAGATCGCTAGTATGTTCTGTGAACACAAATAAATTTATTTA 1215  
 Qy 1140 TCAGGAGCGACTGTGAGATCGCTAGTATGTTCTGTGAACACAAATAAATTTATTTA 1199  
 Db 1216 CTGTC 1220  
 Qy 1200 CTGTC 1204  
 RESULT 3 AF039067 1309 bp mRNA PRI 22-AUG-1998  
 LOCUS Homo sapiens anti-death protein (IEX-1L) mRNA, complete cds.  
 DEFINITION AF039067  
 ACCESSION AF039067  
 NID 9349375  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryotes; Chordata; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1309)  
 AUTHORS Wu, M.X., Ao, Z., Prasad, K.V.S., Wu, R. and Schlossman, S.F.  
 TITLE IEX-1L, an apoptosis inhibitor involved in NF-kappaB-mediated cell survival  
 JOURNAL Science 281 (5379), 998-1001 (1998)  
 MEDLINE 98369175  
 REFERENCE 2 (bases 1 to 1309)  
 AUTHORS Ao, Z. and Wu, M.X.  
 TITLE Direct Submission

JOURNAL Submitted (17-DEC-1997) Tumor Immunology, Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA

FEATURES

source 1. .1309

gene /organism="Homo sapiens"

1. .1309 /db\_xref="taxon:9606"

CDS /gene="IEX-1L"

1. .582 /gene="IEX-1L"

/note="type II membrane protein"

/codon\_start=1

/product="anti-death protein"

/db\_xref="PID:g3449376"

/translation="MCHSRSCPTMTILQAPTAPSTIPGRRSGPEITFDPLPEP  
AAPAGRPASRGRKRRSRVSLPSSGEYRRSGSCALPWSHWGTRLSQSLDLCLL  
ECLPLVRRQLPVEEPNPAKRLLELLTIVFCQILMAEEGVPALPEDAPNAAASLPT  
PVSVPLEPNTSEPSDIALDUSTFLQQHPAA"

BASE COUNT 262 a 393 c 369 g 285 t

ORIGIN

very Match 79.58; Score 976; DB 27; Length 1309;

Best Local Similarity 99.88; Pred. No. 0.00e+00;

Matches 988; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Db 322 GTCCGGCCGAGCTGCCAGTCGAGGAACCGAACCCAGCCAAAGGCTTCTCTTTCTGCTG 381

Qy 211 GTCCGGCCGAGCTGCCAGTCGAGGAACCGAACCCAGCCAAAGGCTTCTCTTTCTGCTG 270

Db 382 CTCACATCGTCTTCTCCAGATCCTGATGGCTGAAGAGGGTGTGCCGGGCCCTGCCT 441

Qy 271 CTCACATCGTCTTCTCCAGATCCTGATGGCTGAAGAGGGTGTGCCGGGCCCTGCCT 330

Db 442 CCAGAGACGCCCTAAGCGGATCCCTGCGCCACCCCTGTGTCCTCCGCTCGAG 501

Qy 331 CCAGAGACGCCCTAAGCGGATCCCTGCGCCACCCCTGTGTCCTCCGCTCGAG 390

Db 502 CCCTTTAATCTGACTTCGGAGCCCTCGGACTACGCTCTGGACCTCAGCACTTTCCTCCAG 561

Qy 391 CCCTTTAATCTGACTTCGGAGCCCTCGGACTACGCTCTGGACCTCAGCACTTTCCTCCAG 450

Db 562 CAACACCGCGCGCTTCTAATCTGACTCCCGCACTCCCAAAAGAAATCGCAAAAC 621

Qy 451 CAACACCGCGCGCTTCTAATCTGACTCCCGCACTCCCAAAAGAAATCGCAAAAC 510

Db 622 CACAAGAAACACAGCGGTACCTGGTCCGAGAGCGGTATCCCACTGGACTTCCGA 681

Qy 511 CACAAGAAACACAGCGGTACCTGGTCCGAGAGCGGTATCCCACTGGACTTCCGA 570

Db 682 GGCAACTTGAACCTCAGAACACTACAGCGGAGACGCCACCGGTGCTTGAGCGGGACCGA 741

Qy 571 GGCAACTTGAACCTCAGAACACTACAGCGGAGACGCCACCGGTGCTTGAGCGGGACCGA 630

Db 742 GGCGCAGAGACCGAGCGGCATAGACCGGAGCAGCCACCGCTGGG-CTAGGCG-GG 799

Qy 631 GGCGCAGAGACCGAGCGGCATAGACCGGAGCAGCCACCGCTGGGCTAGGCGCGCG 690

Db 800 TGGGAAGGAGCGCTCTTAATTTATTCTTATGCTCCTAATAATATTATATGATT 859

Qy 691 TGGGAAGGAGCGCTCTTAATTTATTCTTATGCTCCTAATAATATTATATGATT 750

Db 860 TATGTAGCTCTCTAGGTGATGAGATGTGTAGCTAATATTTATTTAACTTATGCAAG 919

Qy 751 TATGTAGCTCTCTAGGTGATGAGATGTGTAGCTAATATTTATTTAACTTATGCAAG 810

Db 920 GGTGTAGATGTTCCCTCTCTGCTGTAATGCAAGTCTCTTGGTATTTATGAGCTTTTGG 979

Qy 811 GGTGTAGATGTTCCCTCTCTGCTGTAATGCAAGTCTCTTGGTATTTATGAGCTTTTGG 870

Db 980 GACTGTGGAGGAGGACACCTTGAACCTGGCGCAAGTAGGAGAGAAATGGGGAGGACT 1039

Qy 871 GACTGTGGAGGAGGACACCTTGAACCTGGCGCAAGTAGGAGAGAAATGGGGAGGACT 930

Db 1040 CCGGTGGGGAGGAGCGTCCCGCTGGGATGAAGTCTGGTGGTCTGTAAGTTTAGGAG 1099

Qy 931 CCGGTGGGGAGGAGCGTCCCGCTGGGATGAAGTCTGGTGGTCTGTAAGTTTAGGAG 990

Db 1100 GTGACTGCATCTCCAGCATCTCAACTCCGCTCTGCTACTGTGTGAGACTTCGGCGGACC 1159

Qy 991 GTGACTGCATCTCCAGCATCTCAACTCCGCTCTGCTACTGTGTGAGACTTCGGCGGACC 1050

Db 1160 ATTAGGAATGAGATCGTGGATCCCTTCCATCTTCTTGAAGTCGCCCTTTAGGTTGGCTGC 1219

Qy 1051 ATTAGGAATGAGATCGTGGATCCCTTCCATCTTCTTGAAGTCGCCCTTTAGGTTGGCTGC 1110

Db 1220 GAGTAGAGGGTGGGGTGGTGGCTGTCCAGGAGCACTGTCCAGATCGCCTAGTAT 1279

Qy 1111 GAGTAGAGGGTGGGGTGGTGGCTGTCCAGGAGCACTGTCCAGATCGCCTAGTAT 1170

Db 1280 GTTCTGTGAACACAATAAATGATTAT 1309

Qy 1171 GTTCTGTGAACACAATAAATGATTAT 1200

RESULT 4

LOCUS HSPRG1 1864 bp DNA PRI 23-SEP-1997

DEFINITION H.sapiens PRG1 gene.

ACCESSION X96438

NID g2440072

KEYWORDS PRG1 gene.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1864)

AUTHORS Trauzold, A.

TITLE Direct Submission

JOURNAL Submitted (05-MAR-1996) A. Trauzold, Laboratory of Molecular

Gastroenterology, 1st Dept. of Medicine, University of Kiel,

Schittenhelmstrasse 12, Kiel, D-24105, FRG

REMARK revised by [4]

REFERENCE 2 (bases 1 to 508)

AUTHORS Schaefer, H., Trauzold, A., Lettau, P., Kalthoff, H., Foelsch, U.R. and

Schmidt, W.B.

TITLE cDNA cloning and sequencing of a novel human early response gene

and characterization of its expression in pancreatic carcinoma

cells

JOURNAL Gastroenterology In press

REFERENCE 3 (bases 1 to 1864)

AUTHORS Schaefer, H.

TITLE Direct Submission

JOURNAL Submitted (23-SEP-1997) H. Schaefer, Trauzold, Laboratory of

Molecular Gastroenterology, 1st Dept. of Medicine, University of

Kiel, Schittenhelmstrasse 12, Kiel, D-24105, FRG

COMMENT On Sep 27, 1997 this sequence version replaced gi:1515291.

FEATURES

source 1. .1864

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/cell\_line="818-4"

1. .562

/gene="PRG1"

1. .1864

/gene="PRG1"

Join(563..801,914..1864)

/gene="PRG1"

563..801

/gene="PRG1"

/number=1

Join(592..801,914..1174)

/gene="PRG1"

/codon\_start=1

/db\_xref="PID:g350480"

/db\_xref="PID:g2440073"

/translation="MCHSRSCPTMTILQAPTAPSTIPGRRSGPEITFDPLPEP  
AAPAGRPASRGRKRRSRVLYPRVRRQLPVEEPNPAKRLLELLTIVFCQILMAE

	intron	exon	BASE COUNT	395 a	571 c	519 g	379 t			
	EGVPAPLPEDAPNAASLAPTVPVSPVLEPNLTSPEVDALDLSTFLQOHPAAF	802..913 /gene="PRG1" /number=1 914..1864 /gene="PRG1" /number=2								
	Query Match	74.7%;	Score	917;	DB	26;	Length	1864;		
	Best Local Similarity	99.3%;	Pred. No.	0.00e+00;						
	Matches	949;	Conservative	0;	Mismatches	2;	Indels	5;	Gaps	5;
Db	913	GGTCCGCGCCAGCTGCCAGTCGAGAAACCGAACCCAGCCAAAAGGCTTCTCTTTCTGCT	972							
Qy	210	GGTCCGCGCCAGCTGCCAGTCGAGAAACCGAACCCAGCCAAAAGGCTTCTCTTTCTGCT	269							
Db	973	GCTCACATGTCCTTCTGCCAGATCCTGATGGCTGAAGAGGGTGTGCCGGGCCCTGCC	1032							
Qy	270	GCTCACATGTCCTTCTGCCAGATCCTGATGGCTGAAGAGGGTGTGCCGGGCCCTGCC	329							
Db	1033	TCAGAGGAGCCCTTAAGCCGCGCATCCCTGGCGCCACACCTGTGTCCCCCTCCTCGA	1092							
Qy	330	TCAGAGGAGCCCTTAAGCCGCGCATCCCTGGCGCCACACCTGTGTCCCCCTCCTCGA	389							
Db	1093	GCCTTTAATCTGACTTCGGAGCCCTCGGACTAGCTCTGGACCTCAGCACTTTCCTCCA	1152							
Qy	390	GCCTTTAATCTGACTTCGGAGCCCTCGGACTAGCTCTGGACCTCAGCACTTTCCTCCA	449							
Db	1153	GCAACACCCGGCGCCCTTCTAACTG-GACTCCCGCACTCCCCAAAAGAAATCCGAAAA	1211							
Qy	450	GCAACACCCGGCGCCCTTCTAACTGACTCTCCCGCACTCCCCAAAAGAAATCCGAAAA	509							
Db	1212	CCACAAAGAAACACGAGGCGTACCTGGTGTGGCGAGAGCGGTATCCCCAATGGGACTTC	1271							
Qy	510	CCACAAAGAAACACGAGGCGTACCTGGT-GCGCGAGAGCGGTATCCCCAATGGGACTTC	568							
Db	1272	GAGCAACTTGAATCAGAACACTACAGCGGAGAGCGCCACCCGGTGTGTGAG-CGGGACC	1330							
Qy	569	GAGCAACTTGAATCAGAACACTACAGCGGAGAGCGCCACCCGGTGTGTGAGCGGGACC	628							
Db	1331	GAGCGCACAGAGACCGCGCATAGAGCCGAG-CACAGCCACGCTGGG-CTAGGCCCC	1388							
Qy	629	GAGCGCACAGAGACCGCGCATAGAGCCGAGCGCCAGCGCACAGCCAGCTGGGGCTAGGCC	688							
Db	1389	GGTGGGAAGGAGAGCGTCGTTAATTTATTTCTATTGTCTCTAATTAATATTTATGTA	1448							
Qy	689	GGTGGGAAGGAGAGCGTCGTTAATTTATTTCTATTGTCTCTAATTAATATTTATGTA	748							
Db	1449	TTTATGTACGTCCTCTAGTGATGGAGATGTAGCTAAATTTATTTAACTATATGCA	1508							
Qy	749	TTTATGTACGTCCTCTAGTGATGGAGATGTAGCTAAATTTATTTAACTATATGCA	808							
Db	1509	AGGCTGTGAGATGTTCCCTCTGCTGCTAAATTCAGGTCCTCTTGGTATTTATGACCTTGT	1568							
Qy	809	AGGCTGTGAGATGTTCCCTCTGCTGCTAAATTCAGGTCCTCTTGGTATTTATGACCTTGT	868							
Db	1569	GGGTCTGGTGAAGCAGGACACCTGGAAGCTGGCGCAAGTAGGAGCAAGAAATGGGGAGA	1628							
Qy	869	GGGACTGGTGAAGCAGGACACCTGGAAGCTGGCGCAAGTAGGAGCAAGAAATGGGGAGA	928							
Db	1629	CTCGGTGGGGAGGAGCGTCCCGGCTGGGATGAAGTCTGGTGGTGGCTCGTAAGTTAGG	1688							
Qy	929	CTCGGTGGGGAGGAGCGTCCCGGCTGGGATGAAGTCTGGTGGTGGCTCGTAAGTTAGG	988							
Db	1689	AGGTGACTGCATCCTCCAGCATCTCAACTCCGTCGTCTACTGTGTGAGACTTCGGCGGA	1748							
Qy	989	AGGTGACTGCATCCTCCAGCATCTCAACTCCGTCGTCTACTGTGTGAGACTTCGGCGGA	1048							
Db	1749	CAATTAGGAATGAGATCCGTGAGATCCTCCATCTCTTCTTGAAGTCCGCTTATAGGTGGCT	1808							
Qy	1049	CAATTAGGAATGAGATCCGTGAGATCCTCCATCTCTTCTTGAAGTCCGCTTATAGGTGGCT	1108							

Db	1809	CGCAGGTAGAGGGTTGGGGTGTGGGCTCTCAGGAGCGCACTGTGAGATCGCC	1864
Qy	1109	CGCAGGTAGAGGGTTGGGGTGTGGGCTCTCAGGAGCGCACTGTGAGATCGCC	1164

  

RESULT	5	AF083421	477 bp	mRNA	PRI	02-SEP-1998
LOCUS						
DEFINITION		Homo sapiens radiation-inducible immediate early response gene IEX1 (IEX1) mRNA, complete cds.				
ACCESSION		AF083421				
NID		93511288				
KEYWORDS						
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS		1 (bases 1 to 477)				
TITLE		Kondratyev, A.D., Chung, K.N. and Jung, M.O.				
JOURNAL		Identification and characterization of a radiation-inducible				
MEDLINE		glycosylated human early-response gene				
REFERENCE		Cancer Res. 56 (7), 1498-1502 (1996)				
AUTHORS		2 (bases 1 to 477)				
TITLE		Kumar, R., Kobayashi, T., Warner, G.M., Wu, Y., Salisbury, J.L.,				
JOURNAL		Lingle, W. and Pittelkow, M.R.				
AUTHORS		3 (bases 1 to 477)				
TITLE		Kumar, R., Pittelkow, M.R. and Warner, G.M.				
JOURNAL		Direct Submission				
FEATURES		Submitted (11-AUG-1998) Nephrology Research, Mayo Clinic, 200 1st Street SW, Rochester, MN 55905, USA				
source		Location/Qualifiers				
gene		1. .477				
CDS		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		/cell_type="keratinocyte; fetal osteoblast"				
		1. .477				
		/gene="IEX1"				
		7. .477				
		/gene="IEX1"				
		/note="UV irradiation induced gene in keratinocytes"				
		/codon_start=1				
		/product="radiation-inducible immediate early response gene IEX1"				
conflict		/db_xref="PID:g3511289"				
		/translation="MCHSRSCPTMTILQAPTAPSTIPPRGSGPIFTDPLEP				
		AAAPAGPSASGRHKRRRYLYPRVRRQLPVEEPNPKRLFLLLTIVFCQILMA				
		EGVPAPLPEDAPNAASLAPTPVSPVLEPFLNLTSEPSDYALDLSFTLQHPAAF"				
		167				
		/gene="IEX1"				
		/citation=[1]				
		/replace="g"				
		323				
		/gene="IEX1"				
		/citation=[1]				
		/replace="g"				
BASE COUNT		72 a	205 c	117 g	83 t	
ORIGIN						
Query Match		38.4%; Score 471; DB 27; Length 477;				
Best Local Similarity		100.0%; Pred. No. 7,53e-276;				
Matches		471; Conservative	0; Mismatches	0; Indels	0; Gaps	0;
Db	7	ATGTGTCACTCTCGAGCTGCCACCGACCATGACCATCTTCGAGGCCCGCCGCGCC	66			
Qy	1	ATGTGTCACTCTCGAGCTGCCACCGACCATGACCATCTTCGAGGCCCGCCGCGCC	60			
Db	67	CCCTCCACCATCCCGGACCCCGGGGGTCCGGTCTTGAGATCTTCACCTTCGACCCT	126			
Qy	61	CCCTCCACCATCCCGGACCCCGGGGGTCCGGTCTTGAGATCTTCACCTTCGACCCT	120			

```

Db 127 CTCCEGAGCCGCGAGCGCCCTCCCGGGGCCCCAGCGCTCTCGGGGACCCAGAAAG 186
    |||
Qy 121 CTCCEGAGCCGCGAGCGCCCTCCCGGGGCCCCAGCGCTCTCGGGGACCCAGAAAG 180
    |||
Db 187 CGCAGCCGAGGGTCTCTACCTCGAGTGGTTCGCGGCGCCAGCTGCCAGTCGAGGAACG 246
    |||
Qy 181 CGCAGCCGAGGGTCTCTACCTCGAGTGGTTCGCGGCGCCAGCTGCCAGTCGAGGAACG 240
    |||
Db 247 AACCCAGCAAAAGCTTCTCTTTCTGCTGCATCCATCGCTTCTGCCAGATCCTGATG 306
    |||
Qy 241 AACCCAGCAAAAGCTTCTCTTTCTGCTGCATCCATCGCTTCTGCCAGATCCTGATG 300
    |||
Db 307 GCTGAAGAGGGTGTCCCGGCGCCCTGCTCCAGAGGAGCGCCCTTAACGCGCATCCCTG 366
    |||
Qy 301 GCTGAAGAGGGTGTCCCGGCGCCCTGCTCCAGAGGAGCGCCCTTAACGCGCATCCCTG 360
    |||
Db 367 GGGCCACCCCTGTGTCGCCCGCTCTCGAGCGCTTAACTGACTTCGAGGACCCCTCGAC 426
    |||
Qy 361 GGGCCACCCCTGTGTCGCCCGCTCTCGAGCGCTTAACTGACTTCGAGGACCCCTCGAC 420
    |||
427 TACGCTCTGGACCTCAGACATTTCTCCAGCAACACCGCGGCTTCTAA 477
    |||
Qy 421 TACGCTCTGGACCTCAGACATTTCTCCAGCAACACCGCGGCTTCTAA 471
    |||

RESULT 6
LOCUS HSA227914 343 bp mRNA PRI 02-JUN-1998
DEFINITION Homo sapiens partial mRNA; ID YG40-2.
ACCESSION AJ227914
NID 93183967
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 343)
AUTHORS Dominguez,O., Ashhab,Y., Sebater,L., Bellosso,E., Caro,P. and Pujol-Borrelli,R.
TITLE Cloning of labile mRNAs by AU-motif directed display
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 343)
AUTHORS Dominguez,O.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-1998) Dominguez O., Immunologia, Hospital Germans
Trias i Pujol, 08916 Badalona, and Research Center, Almirall
Prodesfarma, 08024 Barcelona, Spain
COMMENT Overlapping sequences: Y14551.
FEATURES
    source
        1..343
            Location/Qualifiers
                1..343
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /dev_stage="adult"
                    /cell_type="T lymphocyte"
                    /cell_line="T122.12"
                    /clone="PH19-20, PH19-21"
                    <1..>343
                    /gene="DIF-2"
                    /gene="DIF-2"
                    /gene="DIF-2"
BASE COUNT 72 a 67 c 114 g 85 t 5 others
ORIGIN
    Query Match 25.5%; Score 313; DB 26; Length 343;
    Best Local Similarity 97.1%; Pred. No. 3.31e-173;
    Matches 333; Conservative 0; Mismatches 7; Indels 3; Gaps 3;

Db 1 GCTTTGTGGACTGTGGAGCAGAC-CCTGGAACTCGCAAGTAGGAGAA-TG 58
    |||
Qy 862 GCTTTGTGGACTGTGGAGCAGACACCTGGAACTCGCAAGTAGGAGAA-TG 921
    |||
Db 59 GGGAGGACTCGGGTGGGAGGAGCAGCTCCCGGCTGGGATGAAGTCTGGTGGGTCGTAA 118
    |||

```

```

Qy 922 GGGAGGACTCGGGTGGGGAGGAGCTCCCGCTGGGATGAAGTCTGCTGGTGGTCTGTAA 981
    |||
Db 119 GTT-AGGAGTGACTGCTCCTCAGCATCTCAACTCGGCTCTGCTACTCTGTGAGACTT 177
    |||
Qy 982 GTT-AGGAGTGACTGCTCCTCAGCATCTCAACTCGGCTCTGCTACTCTGTGAGACTT 1041
    |||
Db 178 CGGCGGACCATTAGGAATGAGATCCGTGAGATCCCTTCCATCTTCTTGAAGTCGCTTTAN 237
    |||
Qy 1042 CGGCGGACCATTAGGAATGAGATCCGTGAGATCCCTTCCATCTTCTTGAAGTCGCTTTAG 1101
    |||
Db 238 GGTGGCTGCGANGTANAGGTGGGGTGGTGGGCTGTCACNAGAGCATATCGACATC 297
    |||
Qy 1102 GGTGGCTGCGANGTANAGGTGGGGTGGTGGGCTGTCACNAGAGCATATCGACATC 1161
    |||
Db 298 GCCTANTATGTTCTGTGAACACAAATAAATTTGATTTACTGTC 340
    |||
Qy 1162 GCCTAGTATGTTCTGTGAACACAAATAAATTTGATTTACTGTC 1204
    |||

RESULT 7
LOCUS MMGLY96 1938 bp RNA ROD 11-MAR-1993
DEFINITION M.musculus gly96 mRNA.
ACCESSION X67644
NID 9287803
KEYWORDS gly96 gene; glycosylated protein.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1938)
AUTHORS Lau,L.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-1992) L. Lau, University of Illinois College of
Med., Dept. of Genetics, M/C 669, 808 South Wood Street, Chicago,
IL 60612, USA
REFERENCE 2 (bases 1 to 1938)
AUTHORS Charles,C.H., Yoon,J.K., Simske,J.S. and Lau,L.F.
TITLE Genomic structure, cDNA sequence, and expression of gly96, a growth
factor-inducible immediate-early gene encoding a short-lived
glycosylated protein
JOURNAL Oncogene 8 (3), 797-801 (1993)
MEDLINE 93173526
FEATURES
    source
        1..1938
            Location/Qualifiers
                /organism="Mus musculus"
                /strain="BALB/c"
                /db_xref="taxon:10090"
BASE COUNT 390 a 525 c 503 g 520 t
ORIGIN
    Query Match 9.6%; Score 118; DB 28; Length 1938;
    Best Local Similarity 77.8%; Pred. No. 6.27e-50;
    Matches 165; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Db 712 ATGTGCCACTCGCGCAACCATCTCCACACCATGACTGGCTGAGGGCCCTTCTCCAGCT 771
    |||
Qy 1 ATGTGTCACTCTCGCAGCTGCCACCCGACCATCTCTGACGAGCCCGGCC 60
    |||
Db 772 CCCTCCACCGCGCCGGAACCTCCGGCGGGGCTCTGCTCCGAAATTTTACGCTTCGAGCCCT 831
    |||
Qy 61 CCCTCCACCATCCGGGAGCCCGCGGGGGTCCGGTCTTGAGATCTTCACCTTCGAGCCCT 120
    |||
Db 832 CTCGCCGAGCGGCGCGGTGTTCACCGCGGTTTGAACACTTCTCGCGGACACCGAAAA 891
    |||
Qy 121 CTCGCCGAGCGCGGAGCGGCGGCTTCGCGGCGGCCAGCGCTCTCGCGGACCCGAAAG 180
    |||
Db 892 CGCAGCCGAGGGTGTCTTACCTTCGAGTGGT 923
    |||
Qy 181 CGCAGCCGAGGGTGTCTTACCTTCGAGTGGT 212
    |||

RESULT 8

```

```

LOCUS      RNPRG1      1758 bp      DNA      03-SEP-1996
DEFINITION R. norvegicus PRG1 gene.
ACCESSION  X96437
NID        g1515318
KEYWORDS   PRG1 gene.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
REFERENCE  1 (bases 1 to 1758)
AUTHORS    Schafer, H., Trauzold, A., Siegel, E.G., Folsch, U.R. and Schmidt, W.E.
TITLE      PRG1: a novel early-response gene transcriptionally induced by
           pituitary adenylate cyclase activating polypeptide in a pancreatic
           carcinoma cell line
JOURNAL    Cancer Res. 56 (11), 2641-2648 (1996)
MEDLINE    96221139
REFERENCE  2 (bases 1 to 1758)
AUTHORS    Trauzold, A.
TITLE      Direct Submission
JOURNAL    Submitted (05-MAR-1996) A. Trauzold, Laboratory of Molecular
           Gastroenterology, 1st Dept. of Medicine, University of Kiel,
           Schittenhelmstrasse 12, Kiel, D-24105, FRG
FEATURES   source
           1..1758
           /organism="Rattus norvegicus"
           /db_xref="taxon:10116"
           /cell_line="AR4-2J"
           590..1279
           /gene="PRG1"
           590..1279
           /gene="PRG1"
BASE COUNT 358 a 475 c 470 g 455 t
ORIGIN
Query Match 9.3%; Score 114; DB 28; Length 1758;
Best Local Similarity 76.9%; Pred. No. 1.67e-47;
Matches 163; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Db 590 ATGTGCTACCTCGGTAGCAGTCCACACCATGCTGCGGCTCCGCGGCCAGTT 649
Qy 1 ATGTGCTACCTCGGTAGCAGTCCACACCATGCTGCGGCTCCGCGGCCAGTT 60
Db 650 CCTCCACGGCCGGAAGTCCGCGGGGCTCCGCTCCCGGAATTTTACCTTCGACCC 709
Qy 61 CCTCCACCATCCCGGAGCCCGCGGGGCTCCGCTCCGATCTTCACCTTCGACCT 120
710 CTCGCGGAGCGGTAGCGTCCGCGCTGCTCCGGAACACTTCTCGCGGCGACGAAA 769
Qy 121 CTCGCGGAGCGGTAGCGTCCGCGCTGCTCCGGAACACTTCTCGCGGCGACGAAA 180
Db 770 CGCAGCGGAGGTGCTCTACCTCGAGTGGT 801
Qy 181 CGCAGCGGAGGTGCTCTACCTCGAGTGGT 212

RESULT 9
LOCUS      I66494      7218 bp      DNA      PAT      23-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION  I66494
NID        92724471
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 7218)
AUTHORS    Dörner, F., Scheiflinger, F. and Falkner, F. Gunter.
TITLE      Recombinant fowlpox virus
JOURNAL    Patent: US 5670367-A 14 23-SEP-1997;
           Location/Qualifiers
           source
           1..7218
           /organism="unknown"
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN
Query Match 5.5%; Score 67; DB 21; Length 7218;
Best Local Similarity 2.4%; Pred. No. 6.44e-20;
Matches 9; Conservative 214; Mismatches 156; Indels 0; Gaps 0;

Db 1063 CGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1122
Cp 1226 CGAGTTTTTTTTTTTTTTTGACAGTAATCAATTTATTTGTTTCACAGACATAC 1167
Db 1123 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1182
Cp 1166 TAGCGCATCTCGAGTCGCTCGTGACAGCCACCAACCCCAACCTCTACCTCGAG 1107
Db 1183 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1242
Cp 1106 CCACCCCTAAAGGCGACTTCAAGAGATGGAAGGATCTCAGGATCTCATCTTAATGCT 1047
Db 1243 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1302
Cp 1046 CGCGAGTCTCACACAGTAGACAGAGGAGTTCAGATGCTGGAGGATGACGATCCTCC 987

```

## ORIGIN

```

Query Match 6.4%; Score 79; DB 21; Length 7218;
Best Local Similarity 1.3%; Pred. No. 9.39e-27;
Matches 5; Conservative 220; Mismatches 146; Indels 0; Gaps 0;

Db 1061 TGGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1120
Qy 99 TGAGATCTTCACTTCGACCTCTCCGAGAGCCGCGAGGCCCTGCGGGGCCCCAG 158
Db 1121 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1180
Qy 159 CGCTCTCGGGCAGCGAAGCGAGCCGAGGTCTCTACCTCGAGTGTCCGGCG 218
Db 1181 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1240
Qy 219 CCAGCTGCCAGTCGAGGAACCGAACCCAGCAGCAAGAGCTTCTCTTCTGCTCACCAT 278
Db 1241 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1300
Qy 279 CGTCTTCTGCCAGATCTGATGCTGAAGAGGTGTGCGGCGGCCCTGCTCCAGAGGA 338
Db 1301 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1360
Qy 339 CGCCCTAACCGCATCCTCGGCGCCACCTGTGCTCCCGCTCTCGAGCCCTTAA 398
Db 1361 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1420
Qy 399 TCTGACTTCGGAGCCCTCGGACTACGCTCTGGACCTCAGCATTCTCTCCAGCAACACC 458
Db 1421 YYYYYYYYYY 1431
Qy 459 GGCGGCTTCT 469

```

## RESULT 10

```

LOCUS      I66494      7218 bp      DNA      PAT      23-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION  I66494
NID        92724471
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 7218)
AUTHORS    Dörner, F., Scheiflinger, F. and Falkner, F. Gunter.
TITLE      Recombinant fowlpox virus
JOURNAL    Patent: US 5670367-A 14 23-SEP-1997;
           Location/Qualifiers
           source
           1..7218
           /organism="unknown"
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN
Query Match 5.5%; Score 67; DB 21; Length 7218;
Best Local Similarity 2.4%; Pred. No. 6.44e-20;
Matches 9; Conservative 214; Mismatches 156; Indels 0; Gaps 0;

Db 1063 CGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1122
Cp 1226 CGAGTTTTTTTTTTTTTTTGACAGTAATCAATTTATTTGTTTCACAGACATAC 1167
Db 1123 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1182
Cp 1166 TAGCGCATCTCGAGTCGCTCGTGACAGCCACCAACCCCAACCTCTACCTCGAG 1107
Db 1183 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1242
Cp 1106 CCACCCCTAAAGGCGACTTCAAGAGATGGAAGGATCTCAGGATCTCATCTTAATGCT 1047
Db 1243 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1302
Cp 1046 CGCGAGTCTCACACAGTAGACAGAGGAGTTCAGATGCTGGAGGATGACGATCCTCC 987

```



AUTHORS  
Pietzsch, A.

## FEATURES

DOE Joint Genome Institute.  
Location/Qualifiers

```

source      1. 74371
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="5q"
/clone="11913"
/chromosome="5"
/note="LBNL H175"
893. 1030
/rpt_family="Alu"
2295. 2438
/rpt_family="Alu"
2818. 2859
/note="(GT)21"
/rpt_type=tandem
/rpt_unit=GT
join(3246. 3410,3721. 3828)
/standard_name="RLF"
/note="65% & 69% protein identity GenPept.U22377"
3431. 3724
/rpt_family="Alu"
3707. 3728
/note="(A)22"
/rpt_type=tandem
/rpt_unit=A
4366. 4661
/rpt_family="Alu"
5327. 5602
/rpt_family="Alu"
6586. 6956
/rpt_family="L1"
6647. 6684
/note="(CA)19"
/rpt_type=tandem
/rpt_unit=CA
7113. 7373
/rpt_family="Alu"
/rpt_type=tandem
/standard_name="possible repeat"
8258. 8503
/rpt_family="Alu"
9070. 9387
/rpt_family="Alu"
complement(9740. 9845)
/rpt_family="MER42"
complement(10440. 11015)
/rpt_family="Alu"
11950. 12250
/rpt_family="Alu"
12057. 12085
/note="(A)29"
/rpt_type=tandem
/rpt_unit=A
12365. 12645
/rpt_family="Alu"
13727. 13750
/note="(AC)12"
/rpt_type=tandem
/rpt_unit=AC
13783. 14024
/rpt_family="L1"
14175. 14470
/rpt_family="Alu"
complement(14906. 15259)
/standard_name="possible repeat"
15300. 15613
/rpt_family="Alu"
16671. 16690
/note="(T)20"
/rpt_type=tandem
/rpt_unit=T
complement(16675. 16977)
/rpt_family="Alu"
complement(16993. 17085)
/rpt_family="MER42"
complement(17678. 18276)
19305. 19583
/rpt_family="Alu"
19914. 19945
/note="(GTT)8"
/rpt_type=tandem
/rpt_unit=GTT
complement(19943. 20222)
/rpt_family="Alu"
21202. 21496
/rpt_family="Alu"
21736. 22035
/rpt_family="Alu"
22017. 22038
/note="(A)22"
/rpt_type=tandem
/rpt_unit=A
22321. 22457
/rpt_family="Alu"
complement(22638. 22981)
/standard_name="possible repeat"
23022. 23326
/rpt_family="Alu"
23473. 23761
/rpt_family="Alu"
23744. 23767
/note="(A)24"
/rpt_type=tandem
/rpt_unit=A
complement(24639. 24694)
/rpt_family="MER42"
complement(25349. 25713)
/rpt_family="Alu"
complement(25727. 26471)
/rpt_family="Alu"
27191. 27477
/rpt_family="Alu"
27774. 28057
/rpt_family="Alu"
28040. 28066
/note="(A)27"
/rpt_type=tandem
/rpt_unit=A
28712. 28930
/rpt_family="MER20"
complement(28769. 28838)
/note="GRAIL 2 excellent exon, frame 0"
28987. 29214
/rpt_family="Alu"
join(29485. 29559,29598. 29779)
/note="95% identity dbSTS:G14522 (SHGC-11312)"
29495. 29976
/standard_name="A1027942"
/note="100% identity EST ov84a10.x1"
complement(30401. 30536)
/note="GRAIL 2 excellent exon, frame 2"
complement(30682. 30733)
/rpt_family="MIR"
complement(31573. 31724)
/note="GRAIL 2 excellent exon, frame 1"
complement(32159. 32232)
/note="GRAIL 2 excellent exon, frame 2"
32388. 32488
/rpt_family="MLT1"
32617. 32908
/rpt_family="Alu"
32977. 33088
/rpt_family="MLT1"
complement(33670. 33785)
/rpt_family="Alu"
complement(34021. 34144)

```

```
misc_feature /note="GRAIL 2 excellent exon, frame 2"
repeat_region /note="GRAIL 2 excellent exon, frame 1"
misc_feature /rpt_family="Alu"
STS 36901..37222
/db_xref="dbEST:AI025011"
/standard_name="STSG-9983"
misc_feature /db_xref="dbSTS:G26554"
misc_feature /note="GRAIL 2 excellent exon, frame 0"
misc_feature complement(join(38063..38218,38462..38578,38741..38995,
39071..39205,39532..39630,39935..40048,40300..40410,
40503..40661,41868..41972,42103..42225,42492..42569,
44379..44507))
/standard_name="histidyl-tRNA synthetase"
misc_feature /note="78%-100% protein identity GenPept:U18937"
misc_feature complement(38089..38215)
misc_feature /note="GRAIL 2 excellent exon, frame 0"
misc_feature complement(38462..38578)
```

...  
Note: remainder of annotations omitted.

```
Query Match 3.1% Score 38; DB 26; Length 74371;
Best Local Similarity 16.7%; Pred. No. 1,38e-04;
Matches 15; Conservative 49; Mismatches 26; Indels 0; Gaps 0;

Db 15957 SSKCWSVYMRMKYCSYCVSGKYYKRCMSYWTYCYKYKWSMXYCTCTSGW 16016
QY 336 GGAGGCCCTAACCGCGATCCCTGGCGCCACCCCTGTGTCGCCGCTCTCGAGCCCT 395
Db 16017 RWMWSKGRMYASRSGCSCSCSMCMCR 16046
QY 396 TAATCTGACTCGGAGCCCTCGGACTACGC 425
```

```
RESULT 14
LOCUS AF012089 10772 bp DNA INV 05-AUG-1997
DEFINITION Drosophila melanogaster cysteine proteinase-1 (CPI) gene, complete
cds, and phenylalanyl tRNA synthetase gene, partial cds.
ACCESSION AF012089
NID 92305220
KEYWORDS fruit fly.
SOURCE Drosophila melanogaster
ORGANISM Eukaryotae; Mitochondrial eukaryotes; Metazoa; Arthropoda;
Tracheata; Hexapoda; Insecta; Pterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 4546 to 4553)
Gray,Y.H., Tanaka,M.M. and Sved,J.A.
p-element-induced recombination in Drosophila melanogaster: hybrid
element insertion
Genetics 144 (4), 1601-1610 (1996)
MEDLINE 97132596
REFERENCE 2 (bases 1 to 10772)
AUTHORS Gray,Y.H.M., Sved,J.A., Preston,C.R. and Engels,W.R.
TITLE Structure of the cysteine proteinase (CPI) gene of Drosophila
melanogaster and associated mutational effects
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 10772)
AUTHORS Gray,Y.H.M., Sved,J.A., Preston,C.R. and Engels,W.R.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1997) School of Biological Sciences, University
of Sydney, Biology A12, Sydney University, NSW 2006, Australia
FEATURES
source Location/Qualifiers
1..10772
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
join(872..1000,2310..2426,6476..6690,6751..7707)
/gene="Cpi"
/product="cysteine protease"
```

```
gene 872..7707
exon /gene="Cpi"
exon 872..1000
/number=1
intron 1001..2309
exon /gene="Cpi"
2310..2426
/number=2
CDS join(2328..2426,6476..6690,6751..7462)
/codon_start=1
/product="cysteine proteinase-1"
/db_xref="PID:g2305221"
/translation="MRTAVLLPLALLAQAQVAFADVYVNEWHHTFKLEHRKNYQDET
EERFLKIFENKHKIAKHNRFAGKVSFKAVNKYADLLHHEFRLMNGFNYYLHK
OLRAADESPKGVTFISPAHVTLPKSYDWRKTGAVTAVKQOCHGSCWAFSSTGALGEO
HFKSGVLVSLSEONLVDGSKYGNNGCGLMDNAFRYIKDNGGIDTKSKYPYEAID
DSCHFNKTVGATDGTGTDIPQDEKKAEMAVATVGPVSAIDASHESQFYSEGVIN
EPOCHQNLHDHGVLVVGGFGTDESGEDYWLKNSWGTGDKGFIKMLRKNENOCGLAS
ASSYPLV"
intron 2427..6475
misc_feature /gene="Cpi"
4546..4553
/gene="Cpi"
/note="Insertion site of P[Casper](50C)"
exon 6476..6690
/number=3
intron 6691..6750
exon /gene="Cpi"
6751..7707
/number=4
mrna join(<8110..9300,9370..>9532)
CDS join(8110..9300,9370..>9532)
/note="potential orf"
/codon_start=1
/product="phenylalanyl tRNA synthetase"
/db_xref="PID:g2305222"
/translation="MLLTAVOGARHMKSTRCLASSAAPAKSPSSPPOLEVSGSTVA
TDGNTVTPKILSYVGANKHLOTDPKLSIIQRIVNYFYGAVRNQNPFLFSYDQNN
PVTVQONFNLLIPADHVSQRKSDCYIIINQHLRAHTAHQVELSGLDNPLVNG
EYVRDEIDSTHYVPFHQADAVRLTKDLFPRIKRYRWDTYFFFPQPPSLEI
SSWTKPNSPATRAVKLMEHEMKHVLKTLKDLFPRIKRYRWDTYFFFPQPPSLEI
YFDNMVLELCCGIMRHEILQKSGVHQSIGAFVGLERLAWLFDIPDLRFSWDS
GFLOSFEKDLHNLKPKYKPSHYPOCTNDLSFWLPQDIEVDAGFSPNDFYDLVRSVAG
DMVEQISLVNKFHFKPTGKSSVCFRIYVRHMERLTQAEVNEIHKQIASASVDSFNVQ
IX"
BASE COUNT 2929 a 2357 c 2282 g 3046 t 158 others
ORIGIN
Query Match 3.0% Score 37; DB 18; Length 10772;
Best Local Similarity 18.1%; Pred. No. 4.16e-04;
Matches 19; Conservative 57; Mismatches 27; Indels 2; Gaps 2;
Db 1716 MKTYRTWMMKMYTSRTTTSAMMMYTWTSTWKYMWAYAMKMMWTRTWAR-MAWASW 1774
QY 706 CGTTAATTTATTTCTTATTCCTCCCTAAATATATATATATATATATATATGTCAGCTCCTCT 765
Db 1775 ARWKWTSAAAYSAWKKMKWYRAMKKTKWMAWKKWKAAW 1819
QY 766 AGGTGATGAGATG-TGTACGTAATATTTATTTTAACTTATGCAA 809
RESULT 15
LOCUS AF012089 10772 bp DNA INV 05-AUG-1997
DEFINITION Drosophila melanogaster cysteine proteinase-1 (CPI) gene, complete
cds, and phenylalanyl tRNA synthetase gene, partial cds.
ACCESSION AF012089
```



```

misc_feature      join(3246..3410,3721..3828)
                  /standard_name="RLF"
repeat_region     /note="55% & 69% protein identity GenPept:U22377"
repeat_region     /rpt_family="Alu"
                  /rpt_type-tandem
                  /rpt_unit-A
repeat_region     4366..4661
                  /rpt_family="Alu"
repeat_region     5327..5602
                  /rpt_family="Alu"
repeat_region     6586..6956
                  /rpt_family="L1"
repeat_region     6647..6684
                  /note="(CA)19"
                  /rpt_type-tandem
                  /rpt_unit-CA
repeat_region     7113..7373
                  /rpt_family="Alu"
repeat_region     complement(7830..8185)
                  /standard_name="possible repeat"
repeat_region     8258..8503
                  /rpt_family="Alu"
repeat_region     9070..9387
                  /rpt_family="Alu"
repeat_region     complement(9740..9845)
                  /rpt_family="MER42"
repeat_region     complement(10440..11015)
                  /rpt_family="Alu"
repeat_region     11950..12250
                  /rpt_family="Alu"
repeat_region     12057..12085
                  /note="(A)29"
                  /rpt_type-tandem
                  /rpt_unit-A
repeat_region     12365..12645
                  /rpt_family="Alu"
repeat_region     13727..13750
                  /note="(AC)12"
                  /rpt_type-tandem
                  /rpt_unit-AC
repeat_region     13783..14024
                  /rpt_family="L1"
repeat_region     14175..14470
                  /rpt_family="Alu"
repeat_region     complement(14906..15259)
                  /standard_name="possible repeat"
repeat_region     15300..15613
                  /rpt_family="Alu"
repeat_region     16671..16690
                  /note="(T)20"
                  /rpt_type-tandem
                  /rpt_unit-T
repeat_region     complement(16675..16977)
                  /rpt_family="Alu"
repeat_region     complement(16993..17085)
                  /rpt_family="MER42"
repeat_region     complement(17678..18276)
                  /rpt_family="Alu"
repeat_region     19305..19583
                  /rpt_family="Alu"
repeat_region     19914..19945
                  /note="(GTTT)8"
                  /rpt_type-tandem
                  /rpt_unit-GTTT
repeat_region     complement(19943..20222)
                  /rpt_family="Alu"
repeat_region     21202..21496
                  /rpt_family="Alu"
repeat_region     21736..22035
                  /rpt_family="Alu"

repeat_region     22017..22038
                  /note="(A)22"
                  /rpt_type-tandem
                  /rpt_unit-A
repeat_region     22321..22457
                  /rpt_family="Alu"
repeat_region     complement(22638..22981)
                  /standard_name="possible repeat"
repeat_region     23022..23326
                  /rpt_family="Alu"
repeat_region     23473..23761
                  /rpt_family="Alu"
repeat_region     23744..23767
                  /note="(A)24"
                  /rpt_type-tandem
                  /rpt_unit-A
repeat_region     complement(24639..24694)
                  /rpt_family="MER42"
repeat_region     complement(25349..25713)
                  /rpt_family="Alu"
repeat_region     complement(25727..26471)
                  /rpt_family="Alu"
repeat_region     27191..27477
                  /rpt_family="Alu"
repeat_region     27774..28057
                  /rpt_family="Alu"
repeat_region     28040..28066
                  /note="(A)27"
                  /rpt_type-tandem
                  /rpt_unit-A
repeat_region     28712..28930
                  /rpt_family="MER20"
misc_feature      complement(28769..28838)
                  /note="GRAIL 2 excellent exon, frame 0"
repeat_region     28987..29214
                  /rpt_family="Alu"
misc_feature      join(29485..29559,29598..29779)
                  /note="95% identity dbSTS:G14522 (SHCG-11312)"
misc_feature      29495..29976
                  /standard_name="AI027942"
                  /note="100% identity EST ov84a10.x1"
misc_feature      complement(30401..30536)
                  /note="GRAIL 2 excellent exon, frame 2"
repeat_region     complement(30682..30733)
                  /rpt_family="MIR"
misc_feature      complement(31573..31724)
                  /note="GRAIL 2 excellent exon, frame 1"
misc_feature      complement(32159..32232)
                  /note="GRAIL 2 excellent exon, frame 2"
repeat_region     32388..32488
                  /rpt_family="MLT1"
repeat_region     32617..32908
                  /rpt_family="Alu"
repeat_region     32977..33088
                  /rpt_family="MLT1"
repeat_region     complement(33670..33785)
                  /rpt_family="Alu"
misc_feature      complement(34021..34144)
                  /note="GRAIL 2 excellent exon, frame 2"
misc_feature      complement(35238..35331)
                  /note="GRAIL 2 excellent exon, frame 1"
repeat_region     36392..36663
                  /rpt_family="Alu"
misc_feature      36901..37222
                  /note="100% identity EST ou55c09.x1"
                  /db_xref="dbEST:AI025011"
                  /standard_name="STSG-9983"
                  /db_xref="dbSTS:G26554"
                  /note="GRAIL 2 excellent exon, frame 0"
                  /note="GRAIL 2 excellent exon, frame 1"
                  /complement(join(38063..38218,38462..38578,38741..38995,
                  39071..39205,39532..39630,39935..40048,40300..40410,

```

FEATURES	Location/Qualifiers
source	1..192
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
STS	36..177
primer_bind	36..58
primer_bind	complement(155..177)
BASE COUNT	40 a 32 c 69 g 51 t

```

STS
primer
primer
BASE COUNT
ORIGIN

```



```
Query Match      2.8%; Score 34; DB 30; Length 408;
Best Local Similarity 88.6%; Pred. No. 1.06e-02;
Matches 39; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 84 TGCAGATGCTATGCTGGGTGGGAAGTGCAGAGTGACACAT 127
Cp 44 TGCAGATGCTATGCTGGGTGGCAGTGCAGAGTGACACAT 1

RESULT 19
LOCUS      I28278      215 bp      DNA      PAT      30-OCT-1996
DEFINITION Sequence 5 from patent US 5569830.
ACCESSION  I28278
NID         91819054
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 215)
AUTHORS     Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
TITLE       Plant inhibitors of fungal polygalacturonases and their use to
            control fungal disease
JOURNAL     Patent: US 5569830-A 5 29-OCT-1996;
FEATURES    Location/Qualifiers
            source          1..215
            BASE COUNT     15 a      8 c      25 g      26 t      141 others
            ORIGIN          /organism="unknown"

Query Match      2.7%; Score 33; DB 21; Length 215;
Best Local Similarity 12.2%; Pred. No. 3.04e-02;
Matches 25; Conservative 79; Mismatches 101; Indels 0; Gaps 0;

Db 11 VSRATSCNDKAKKDGNTSSWTTDCNRTWGVCDTDTTYRVNDSGHKNKSSANNVYGG 70
Qy 799 AACTTATGCAAGGTGTGAGATGCTCCCTGCTGTAATGTCAGGTCTCTTGATTTAT 858
Db 71 NNVGAAKTHYTHNVSGADSKVTVDYNASGTSSENGTGDNRSGADSVGSSKTAWTSR 130
Qy 859 TGAGCTTTGGGAGCTGGTGGAGCAGGACACCTGGAACCTGGCAAGTAGGAGAAGAA 918
Db 131 NRTGKTANNAVDNRNMGDSVSGDKNTKKHAKNSADKVGSKNNGDRNNRYGTGTSNVS 190
Qy 919 ATGGGAGGAGCTGGGTGGGGAGGAGCTCCCGCTGGGATGAAGTCTGCTGGTGGTGC 978
Db 191 NNCGGGNKRDVSYANNKCCGSCT 215
Qv 979 TAAGTTTAGGAGGTGACTGCATCCT 1003

RESULT 20
LOCUS      HUMRNP70A      1662 bp      mRNA      PRI      21-OCT-1993
DEFINITION Human U1 small nuclear ribonucleoprotein 70 kd protein mRNA,
            complete cds.
ACCESSION  M22636
NID         g337446
KEYWORDS    Ribonucleoprotein; small nuclear ribonucleoprotein.
SOURCE      Human placenta and endothelium, cDNA to mRNA.
ORGANISM    Homo sapiens
            Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Carnivora; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1662)
AUTHORS     Query,C.C., Bentley,R.C. and Keene,J.D.
JOURNAL     Unpublished (1989)
REFERENCE   2 (sites)
AUTHORS     Query,C.C., Bentley,R.C. and Keene,J.D.
TITLE       A common RNA recognition motif identified within a defined U1 RNA
            binding domain of the 70K U1 snRNP protein
JOURNAL     Cell 57 (1), 89-101 (1989)
MEDLINE     89195226
COMMENT     [Cell 57, 89-101 (1989)] sites.
            Draft entry and computer-readable sequence [1] kindly submitted by

FEATURES
source
Location/Qualifiers
1..1662
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 49..81
/notes="type C retrovirus p30gag-like region"
172..1630
/partial
/notes="small ribonucleoprotein 70 kd protein mRNA"
172..1485
/notes="small ribonucleoprotein 70 kd protein"
/codon_start=1
/db_xref="PID:g337447"
/translation="MTQFLPPLNLLAFAPRDPIPLPPLKPLEKPHKHHNOPYCGIAPY
IRFEDPRDAPPTRAETREERERKRIERROQEVETELKMDPHNDPNAQGDAP
KTLFVAVNYDTTESKLRFEEYVGPRIKRIHMYVSKRSGKPGYAEIYEHEDMHS
YKHADGKKIDRRVLDVVRGTVKGRRLGGGLGCTGTRGGADVNIRSGRDTISR
YDERPGSPPLPHRDORRERERERERERERERERERERERERERERERERERER
SRERSKDKDRKRSSRSRERARERERERERERERERERERERERERERERERER
LPGDPGDPGPEKGRDRRERERERERERERERERERERERERERERERERERER
RGGGGGDNGLGGLGDRSDRMVSEGGDGLAPENGLYLMEEAPE"
BASE COUNT      377 a      453 c      587 g      245 t
ORIGIN           Unreported.

Query Match      2.5%; Score 31; DB 25; Length 1662;
Best Local Similarity 89.7%; Pred. No. 2.39e-01;
Matches 35; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 1624 AAATAAATAATTCTCTGTTGCAAAAAAATAAAAAA 1662
Qy 1184 AAATAAATAATTCTCTGTTGCAAAAAAATAAAAAA 1222

RESULT 21
LOCUS      AF050666      3436 bp      mRNA      ROD      26-MAR-1998
DEFINITION Mus musculus glycosylphosphatidylinositol-specific phospholipase D
            precursor (GpId1) mRNA, complete cds.
ACCESSION  AF050666
NID         g2984690
KEYWORDS    house mouse.
SOURCE      Mus musculus
ORGANISM    Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 3436)
AUTHORS     LeBoeuf,R.C., Caldwell,M., Guo,Y., Metz,C., Davitz,M.A., Olson,L.K.
            and Deeg,M.A.
TITLE       cDNA isolation, expression and chromosomal location of the mouse
            glycosylphosphatidylinositol-specific phospholipase D gene (GpId1)
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 3436)
AUTHORS     LeBoeuf,R.C., Caldwell,M., Guo,Y., Metz,C., Davitz,M.A., Olson,L.K.
            and Deeg,M.A.
TITLE       Direct Submission
JOURNAL     Submitted (24-FEB-1998) Pathobiology, Univ. of Washington, Room 305
            Raitt Hall/Box 353410, Seattle, WA 98195, USA
FEATURES    Location/Qualifiers
            source          1..3436
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /chromosome="13"
            /map="proximal"
            /cell_line="alpha-TC6"
            /tissue_type="glucagonoma"
            1..3436
            /gene="GpId1"
            271..2784
            /gene="GpId1"
            /function="degrades GPI anchor of cell-surface mammalian
            proteins in the presence of detergent"
            /note="GPI-PLD"
            /codon_start=1
```

Query Match	2.4%	Score 29	DB 19	Length 1125
Best Local Similarity	89.2%	Pred. No. 1.76e+00		
Matches	33	Conservative	0	Mismatches 4
			0	Indels 0
			0	Gaps 0

  

Db	1089	AATAAAATTATAACTGCCAAAAA	AAAAAAAAA	1125
Qy	1185	AATAAAATTGATTACTGTCAAAAA	AAAAAAAAA	1221

  

RESULT	24			25-JAN-1993
LOCUS	BOVIOPP	1266 bp	mRNA	
DEFINITION	Bovine inorganic pyrophosphatase mRNA sequence.			
ACCESSION	M95283			
NID	g163228			
KEYWORDS	inorganic pyrophosphatase.			
SOURCE	Bos taurus retina cDNA to mRNA.			

```

ORGANISM      Bos taurus
Eukaryoteae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.
REFERENCE     1 (bases 1 to 1266)
AUTHORS      Yang, Z. and Wensel, T.G.
TITLE         Molecular cloning and functional expression of cDNA encoding a
              mammalian inorganic pyrophosphatase
JOURNAL       J. Biol. Chem. 267, 24641-24647 (1992)
MEDLINE       93077559
FEATURES      Location/Qualifiers
               1..1266
               /organism="Bos taurus"
               /db_xref="taxon:9913"
               /tissue_type="retina"
BASE COUNT    396 a      239 c      300 g      331 t
ORIGIN
Query Match   2.4%; Score 29; DB 19; Length 1266;
Best Local Similarity 87.2%; Pred. No. 1.76e+00;
tches 34; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

UD	1227	AAATATAAATATATTTGTTTCAAAAAAAAAAAAAAAAAA	1265
QY	1184	AAATATAAATGATTACTGTCAAAAAAAAAAAAAAAAAA	1222

RESULT	25	DIRHSP70	1294 bp	mRNA	INV	25-AUG-1992
LOCUS		Dirofilaria immitis	70 kda	heat shock protein homologue (hsp 70)		
DEFINITION		mRNA, 3' end.				
ACCESSION		M95648				
NID		g156705				
KEYWORDS		heat shock protein 70; heat shock protein homologue.				
SOURCE		Dirofilaria immitis adult cDNA to mRNA.				
ORGANISM		Dirofilaria immitis				
		Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;				
		Secernentea; Spiruria; Spirurida; Spirurina; Filarioidae;				
		Onchocercidae; Dirofilaria.				
REFERENCE		1 (bases 1 to 1294)				
AUTHORS		Culpepper, J.A., Friedman, L. and Dale, B.				
TITLE		Molecular cloning and characterization of a Dirofilaria immitis cDNA encoding an hsp 70 homologue				
JOURNAL		Unpublished (1992)				
FEATURES		Location/Qualifiers				
source		1. .1294				
		/organism="Dirofilaria immitis"				
		/db_xref="taxon:6287"				
		/dev_stage="adult"				

```

CDS
1..1038
/dev_stage="adult"
/partial
/note="homologue"
/codon_start=1
/product="heat shock protein 70"
/db_xref="PID:g156706"
/translation="REELCADLFRSTMPVEKALDKMKDKAQVHDIIVLVGGSTRIP
KVQKLSDFSGELNKSINPDCAVYGAQVAILSGDKSEAVQDLLLDVAPLSLG
IETAGWQMALIKRIITPTKTSQITFYSDNQPGVLIQVIEGERAMTKDNNLLGFRF
LSGIPAPRGVPOIETFYDANGILNVSAQDKSTQKNTITTDKRLSGRLDEIRM
VQEAERYKADDDAKQIAAKNALESAYFNMKQTIIEDEKLKDKISDDKKKIKQEKDE
TVRWLNDGNTAENKDEFEHRQKELESVCNPIITKLVSAGMGPMGPGMPGGAPGGS
TGGGPRIEEVD"
BASE COUNT      409 a      229 c      296 g      360 t
ORIGIN
Query Match      2.48; Score 29; DB 18; Length 1294;
Best Local Similarity 87.2%; Pred. No. 1.76e+00;
Matches 34; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

Db	1256	AAGTAAATTTATTATCGGGTCAAAAA	1294
Qy	1184	AAATAAATTGATTTACTGTCAAAAAA	1222

RESULT	26	
LOCUS	MM26SPROT	1479 bp mRNA
DEFINITION	Mus musculus mRNA for 26S proteasome non-ATPase subunit.	10-SEP-1998
ACCESSION	Y13071	
NID	q2505939	
KEYWORDS	26S proteasome; non-ATPase subunit; proteolysis; regulatory complex.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	1 (bases 1 to 1479) Penney,M., Wilkinson,C., Wallace,M., Javerzat,J.P., Ferrell,K., Seger,M., Dubiel,W., McKay,S., Allshire,R. and Gordon.C.	
TITLE	The padi1(+) gene encodes a subunit of the 26 S proteasome in fission yeast	
JOURNAL	J. Biol. Chem. 273 (37), 23938-23945 (1998)	
MEDLINE	98395111	
REFERENCE	2 (bases 1 to 1479) Gordon,C.	
AUTHORS	Direct Submission	
TITLE	Submitted (08-MAY-1997) C. Gordon, MRC Human Genetics Unit, Developmental Genetics, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK	
JOURNAL		
FEATURES	Location/Qualifiers	
source	1..1479 /organism="Mus musculus" /db_xref="taxon:10090" /dev_stage="adult" 237..1166 /gene="PADI" 237..1166 /gene="PADI" /function="proteolysis" /codon_start=1 /product="26S proteasome, non-ATPase subunit" /db_xref="PID:e352083" /db_xref="PID:g2505940" /db_xref="SPREMBL:O35593" /translation="MDRLRLGMPGLARPPTDAPVDAVEQVIISLALLKMLKHG RAGVPMEVGLMGLGEFVDYTVRVIDVFAMPSGTGVSVEAVDPVFOAKMLDKMTG LPWVGVWGHSHPGFCWLSGVNDITQSFSEALSERAVAVVDPIQSVKGKVVVIDAF RIANNVVLGHEPROTSLNGLHFKPSIOALLHGLNRHYSLITINYRNELEQKMLLN LHKXSMESGILQDYSEKCHKHNSVYKMELEAKKNYKKAVEEDKMTPEQLAKNNVGK QDPKRHLHEHDVDTNSNIVQCLAAMLDIVFRK"	
gene	453 a 270 c 349 g 407 t	
CDS		
BASE COUNT	453 a 270 c 349 g 407 t	
ORIGIN		
Best Match	2.4%; Score 29; DB 28; Length 1479;	
Query Local Similarity	87.2%; Pred. No. 1.76e+00;	
Matches	34; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	
Db	1440 AAATAAATACATTACATCCCAAAAAAAAAAAAAAAAAAAAA 1478 	
Qy	1184 AAATAAATGATTACTGTCAAAAAAAAAAAAAAAAAAAAA 1222 	
RESULT	27	
LOCUS	XLU69669	1811 bp mRNA
DEFINITION	Xenopus laevis nuclear pore complex-associated protein TPR (tpr mRNA, partial cds.	26-FEB-1997
ACCESSION	U69669	
NID	g1850343	
KEYWORDS	African clawed frog.	
ORGANISM	Xenopus laevis	
REFERENCE	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; pipidae; Xenopodinae; Xenopus.	
AUTHORS	1 (bases 1 to 1811) Cordes,V.C., Reidenbach,S., Rackwitz,H.R. and Franke,W.W.	
TITLE	Identification of protein p270/Tpr as a constitutive component of	

```

/gene="fpa2"
Join(1061..1157,1312..1703)
/gene="fpa2"
/codon="skpl-like; similar to the product of the fpa1 gene"
/codon_start=1
/product="cytosolic glycoprotein FP21"
/db_xref="PID:gl658024"
/translation="MSLVKSSDEKVFIEIKACMSVTIKNMIEDIGSDAIPILP
NYTSTILKLVYCRHHQHSPQGDKKDKRLDDIPYDRDFCKVDQPTLFELILA
ANYLDIKPLDVTCKTVANNRGKTPPEIRKIFNIKNDFTPEEEQIRKENCEWKG
GN"
intron      1158..1311
BASE COUNT  742 a 166 c 170 g 762 t
ORIGIN

Query Match      2.4% Score 29; DB 18; Length 1840;
Best Local Similarity 87.2%; Pred. No. 1.76e+00;
Matches 34; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 915 TTTTNTTTTTTTTTTAAAAATAAATAATTTTATTTT 953
      | | | | | | | | | | | | | | | | | | | |
Cp 1222 TTTTNTTTTTTTTTTTGACAGTAATCAATTTTATTT 1184

RESULT 29
LOCUS      HSMIGGE      1941 bp      DNA      PRI      05-NOV-1995
DEFINITION H. sapiens mig-5 gene.
ACCESSION  Z30183
NID        9520931
KEYWORDS   mig gene.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryote; Eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
AUTHORS    Wick, M., Burger, C., Brüsselbach, S., Lucibello, F.C. and Muller, R.
TITLE       A novel member of human tissue inhibitor of metalloproteinases
            (TIMP) gene family is regulated during G1 progression, mitogenic
            stimulation, differentiation, and senescence
JOURNAL     J. Biol. Chem. 269 (29), 18953-18960 (1994)
MEDLINE     94308155
REFERENCE   Mueller, R.
AUTHORS     Direct Submission
TITLE       Submitted (14-FEB-1994) Rolf Mueller, Institut fuer
            Molekularbiologie und Tumorforschung (IMT), Emil-Mannkopff-Strasse
            2, Marburg, 35037, Germany
FEATURES
source      1..1941
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="30-13 (mig-5)"
            /cell_type="Fibroblast"
            /cell_line="WI-38"
            42..674
            /gene="mig-5"
            42..674
            /gene="mig-5"
            /codon_start=1
            /db_xref="PID:g520932"
            /db_xref="SWISS-PROT:P35625"
            /translation="MTPWGLIVLLGWSMGTCAPCTCSPHQDAFCNSDIVIRAK
            VVKKLVKSGPCTGLVYTIKQMKYRGFTKPMHVOYIHTFEASESLGGLKLEVNKYOL
            LGRGVYDQKGLNFTVFRVYEQQLTISQRKLNRYHLGNCNKIKSCYYILPCFVTSKN
            ECLWTDMLSNFYGGVQSKHYACIRQGGYCSYRWGAPPKSIINATDP-
            3 others
BASE COUNT  506 a 496 c 437 g 499 t
ORIGIN

Query Match      2.4% Score 29; DB 25; Length 1941;
Best Local Similarity 85.4%; Pred. No. 1.76e+00;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```



```
/tissue_type="liver"
/dev_stage="adult"
39..2846
CDS
/codon_start=1
/product="inter-alpha-inhibitor heavy-chain H2"
/db_xref="PID:e307038"
/db_xref="PID:g1915954"
/translation="MGKLTGFLLCFLLESEAGFEIPTNGLSFEAEYCDLAEALGKGF
VFGNRRSGGVQDVLTISYKQSTIISRMANTVIQTKVNHNSPEPDQVDFDIQPKG
AFISNFSMTVDGTFKTSIKETKTVGRALYQWAKGKTAGLVRSRALDMENFKIEVNI
APGAKVQFELHYQEVNKLGSYEHRIHLQGLRAKLAHLEVDVQIIEPQGLRFLHVLDT
FDGHLFGVPPVQVQKQAHVAFPTVAQKRCSPGSAVDGELVVMYDNNRQKAGE
LQFNGVFVEHFAPESDPIKNIILFVIDYSGSMGTGKMKOTVEAKTITLDDRAEQ
FSLVDFNNHRTWRNDLVSAKTKQVADAKTYIEKIOPSGGTNINEALLRAIFILNEAN
NLGLDDPNSVLILVSDGPTGELQSLKIQNVKQNDQNVSLSLGSGFDVDVDF
LKLNSDNRGMAQRIYQNDTASOLKFKFYNOVSTPLLRNVQFNPQASVTDVTONSFP
NYFGGSIVVAGFNPEKLEQAGIITATSANVELVLETLAEQDGLAEFLAKDRHADP
DFTKLLAYVITINOLDESRAPSAVKKITSILQSLDHDHIVPLPLAMVVENAG
DERMLADPQDQSCSTNLNKGKVTNPSLPSWVNPPTPRVPLPAVGPSVIEATPP
PHMREVDHPHFIIYLPSONICFNIDSEPKILANLVSDPESGIVNGOLISAKKIL
DGKLTGFKIGIFQHEDEVKVELISTETISLSRGRSVLSWSDSALYLNQRVHISVK
KEKTVTLVDOEVSFVLLHRVKKHPIVNDVGLIYIPPTKFSKRAHGLIGQFMHEP
EIRIFNRPQKDEPEASMEVKGQTLVTVTGLQKDYRTDRVFGTDVPCWVFVHNSKG
FIDGHYKDYLVPLLYSLKRP"
BASE COUNT      836 a 770 c 779 g 567 t
ORIGIN

Query Match      2.4%; Score 29; DB 19; Length 3042;
Best Local Similarity 78.4%; Pred. No. 1.76e+00;
Matches 40; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Db 2981 TTCCTGTATAATACATTTTTCCTCTGTAAAAAAGAAAAAAGAAAAA 3031
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1172 TTCCTGACACAAATAAATGATTCTGTCACAAAAAAGAAAAAAGAAAAA 1222

RESULT 33
LOCUS      ZMU85494      3217 bp      mRNA      PLN      26-MAY-1998
DEFINITION Zea mays LON1 protease (LON1) mRNA, complete cds.
ACCESSION      U85494
NID      g1816385
KEYWORDS
SOURCE      Zea mays.
ORGANISM      Zea mays.
Eukaryota; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta; Liliopsida;
Poales; Poaceae; Zea.
RENCES 1 (bases 1 to 3217)
JTHORS Barakat,S., Pearce,D.A., Sherman,F. and Rapp,W.D.
TITLE      Maize contains a Lon protease gene that can partially complement a
yeast pimi-deletion mutant
JOURNAL      Plant Mol. Biol. 37 (1), 141-154 (1998)
MEDLINE      98281582
REFERENCE 2 (bases 1 to 3217)
AUTHORS      Rapp,W.D. and Barakat,S.
TITLE      Direct Submission
JOURNAL      Submitted (14-JAN-1997) Biology, University of Missouri-St. Louis,
8001 Natural Bridge Road, St. Louis, MO 63121, USA
FEATURES
source
1..3217
/organism="Zea mays"
/strain="B73"
/db_xref="taxon:4577"
1..3217
/gene="LON1"
/gene="LON1"
/feature="Lon protease homolog"
/codon_start=1
/product="LON1 protease"
/db_xref="PID:g1816386"
/translation="MSDSPVLPRLAVLPFRNKVLLPGLAIVIRICTNPSSVILVEOE
LWQNEKGLIGLVPVRDSEATAVGSLSPGVGSDSGEGSGVSGSAVSESSKQDTKNK
EPIHWSKGAARALHLSRGVEKPSGRVTVIWLGLCRFSVOELSAKGPYHVARVSR
LDMTKTELEQEDPDILALSRQKATAMELISVLEQKQTKGRTKVLLDTPVTVRLA
DIFVAFSEISFEQELSDMLSVHLKATKATLVDHRLQSLIYAEKIKQVEGQLSKS
QKFEFLRQOMRAIKEDDDDDVAALRKMQMGAMPANIMWHAQMRRLRKMQ
POQPGYSSRAYLELLADLPQKVSSEERELDUVAKESSLDQDHYGLTKVKORIIETLA
VRKLKPDARGPVLFCVGPVPGVGTSLASSIAKALNKFTIRISLGGVKDEADIRHRT
YIGSPGRLIDGLKRVSVSNPMLDEIDKGTSDVGPDPASALLLEVLPQONKAFNDH
YLVNPFDSLKVFATANRMPPIPPLLDQMEIIEIPGVTPEKLIKAMKHILPVLE
OHLSTTNLOIPEANVKLVIEREAGVRNLERLALAAARAAVKAQVKTBLGK
EQDAPITTLDSRLADGGVEVEVIMPEHDIISNTENSPMIVDEAMLEKVLGPRFD
DREARADNLSPSTDNILLESRDIIHFPAAGVPGKDPGAGVLTVALVSLFSNRK
VWRAADTGMTLRLGLVPGGVKDKVLAHAHRYGIRKRVILPERNLKDLSEVPLPILSD
MEILLAVKRIEELVDAFEGRCPLRSKSL"
BASE COUNT      881 a 650 c 845 g 841 t
ORIGIN

Query Match      2.4%; Score 29; DB 24; Length 3217;
Best Local Similarity 96.8%; Pred. No. 1.76e+00;
Matches 30; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 3187 TACTTTCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3217
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1198 TACTGTCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1228

RESULT 34
LOCUS      S78355      3737 bp      mRNA      ROD      21-SEP-1995
DEFINITION Cyl-1-cyclin D1 [mice, BALB/c, brain, mRNA, 3737 nt].
ACCESSION      S78355
NID      g994896
KEYWORDS      Mus sp. brain BALB/c.
SOURCE      Mus sp.
Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE 1 (bases 1 to 3737)
JTHORS Smith,R., Peters,G. and Dickson,C.
TITLE      Genomic organization of the mouse cyclin D1 gene (Cyl-1)
JOURNAL      Genomics 25 (1), 85-92 (1995)
MEDLINE      95293413
REMARK      GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 167715] from the original journal article.
This sequence comes from Fig. 1.
Map location: 7.
FEATURES
Location/Qualifiers
1..3737
/organism="Mus sp."
/db_xref="taxon:10095"
138..1025
/note="cyclin D1"
/gene="Cyl-1"
138..1025
/gene="Cyl-1"
/note="This sequence comes from Fig. 1."
/codon_start=1
/product="cyclin D1"
/db_xref="PID:g994897"
/translation="MEHQLLCCCEVETIRRAYPDTNLLNDRVLRAMUKTETCAPSVSY
FKCQKQIVPSNRKIVATWMLVECEQKCEEVFFPLANNVLDLRFSLKPKSRQLL
GATCFVASKMKETIPLTAELKICIVTDSIRPEELQMLLKNLLKWLAAWMTPHDF
IEHFLSKPEADENKQTIKHAQTFVACDVKFISNPPSVAAAGSVAAAGMOGLNG
SPNNELSCYTHHLSRVIKCDPLCRACQIEALLESRLRQAQNVDPKATEEGE
VEEAGLACTPTDVRVDI"
BASE COUNT      921 a 904 c 988 g 924 t
ORIGIN

Query Match      2.4%; Score 29; DB 28; Length 3737;
Best Local Similarity 87.2%; Pred. No. 1.76e+00;
Matches 34; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 3697 AAATAAAATTTGGTAAAAATCTCAAAAAAAGAAAAAAGAAAAA 3735
```



```

||||| 1184 AAATAAATGATTACTGTCAAAAAAAAAAAAAAAAAAAAA 1222
||||| 1184 AAATAAATGATTACTGTCAAAAAAAAAAAAAAAAAAAAA 1222

RESULT 35
LOCUS XELMYBRP1 3789 bp mRNA 11-SEP-1992
DEFINITION Xenopus laevis myb-related protein 1 (myb1) mRNA, complete cds.
ACCESSION M75870
NID 9214597
KEYWORDS myb-related protein 1.
SOURCE Xenopus laevis gastrula, oocyte cDNA to mRNA.
ORGANISM Xenopus laevis
Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea;
Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 3789)
AUTHORS Boumeester, T., Guehmann, S., El-Baradi, T., Kalkbrenner, F., Van
Wijk, D., Moelling, K. and Pieler, T.
TITLE Molecular cloning, expression and in vitro functional
characterization of Myb-related proteins in Xenopus
Mech. Dev. 37, 57-68 (1992)
JOURNAL 92297434
FEATURES
LOCATION/Qualifiers
SOURCE 1..3789
/organism="Xenopus laevis"
/dev_stage="gastrula, oocyte"
61..2262
/gene="myb1"
61..2262
/gene="myb1"
/product="myb-related protein 1"
/db_xref="PID:9214598"
/translation="MSRRSGDLDLQYDSDVPEPKENRVKVKWTPPEDETLKA
LVKHHQGEWKTIASLNRRTEQQOHLRLVLPDLVKGPMTKEDEKVLKLYKYG
TKHTLIAKLRGKQCRERWNLNPEVKLSWTEEDRIICQAKHVLGNRAEII
AKLLPGRTDNVKNHNSITIKRVETGGLVTKASGQOEDREDSYQAAEDQNVHLLS
EPVRSANIPPEPSNLSKLLTKSPGIRSEQSGEGSESATVDSAPKWMVE
YVNLVPGSDIMESDPEAKCLESFDELGEDSTVDGSPTHAAVTDKQPSNVTEYRL
DGHLSLDCNKGELIPSPQAFGTPSPVLQHKRKITLSPVTEGSGSTTSV
TEAMSPTKSPKVSLEPSQOFLNFSKQDALENPSTSTVCSQKOTVTPPLHR
DKPLLOKNSVIFIPNNKFAADHVLHTPTPKNALEKFGSLKPLPPHLEEDLEVL
RSEGIELIIVDEPKRQKPKHSPMKKSLALDIIDKPKPSSLTLPASVSAHM
QPOTCDLSVLSNESCSEENSVLNGFQVKTSGAVVQLGNTSOLLIDIGELYK
TQCSLRNTATPKTNGTINTDLCPQSLMDLDTFHTSTAVGSKNCFITATILFQV
KPEIKMORHMIPEPTAAWKTAVFGSGDQOMLFLSCTVALLILRF5"
1 others
BASE COUNT 1129 a 865 c 837 g 957 t
IGIN

Query Match 2.4%; Score 30; DB 20; Length 3789;
Best Local Similarity 79.7%; Pred. No. 6.55e-01;
Matches 47; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

Db 2989 TTTTITTTTTTTTTCACAGCAAAAAATTTATT-GTGTCAAAAAATAAATAG 3046
||||| 1184 AAATAAATGATTACTGTCAAAAAAAAAAAAAAAAAAAAA 1222
Cp 1222 TTTTITTTTTTTTTCACAGCAAAAAATTTATTGTGTTCACACACATAC 1164

RESULT 36
LOCUS AF064104 4624 bp mRNA PRI 19-MAY-1998
DEFINITION Homo sapiens Cdc14B1 phosphatase mRNA, complete cds.
ACCESSION AF064104
NID 93136331
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 4624)
AUTHORS Hao, L., Baskerville, C. and Charbonneau, H.
TITLE Direct Submission
JOURNAL Submitted (08-MAY-1998) Biochemistry, Purdue University, 1153

FEATURES
LOCATION/Qualifiers
SOURCE 1..4624
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="placenta"
453..1349
/note="CDC14B; similar to Saccharomyces cerevisiae Cdc14p;
similar to the product encoded by GenBank Accession Number
AF023158; putative alternatively spliced variant encoded
by GenBank Accession Number AF064105"
/codon_start=1
/product="Cdc14B1 phosphatase"
/db_xref="PID:93136332"
/translation="MKRSRRSSWAAAPPCSRRCSTSPGVKKIRSTQDPPRRDDP
QDDVLDITDLCLFALYLRKPSASNVHYFSDNELEYENFYADFGLNLAWVRYCC
KINKLKSITMLRKIVHTGTDQKQKAAAFVLCGVMIYVILGRPEAYRILIFGET
SYIPFRAAYSCNFYITLLDCFHAVKAMOVGLFNFSNLDVEYHVEKADENGLMW
IIPDRFTAFGPHSRARLESYHOHSPETYIOYFKNHNTVTLIRLNKMYDAKRETD
GFDHDLFFADGSTPDIAVKEFDICENASGALAVHCKAGLGTGTLIACYIMKHYR
MTAAETIAWRICRPGSVIGQQFLVMKQTNLMLEGDFKQKLGQNGQHRAAFSK
LLSGVDDISINGENQOQPEPEYSDDEINGVTQGDRLALKRRQSKTNAIPLTVI
LQSSVCKTSEPNISSAGITTKRTTSASRKS SVKLSISRTKTCLR"

BASE COUNT 1255 a 1067 c 1115 g 1187 t
IGIN

Query Match 2.4%; Score 29; DB 26; Length 4624;
Best Local Similarity 85.4%; Pred. No. 1.76e-00;
Matches 35; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 3066 ACAGAAACAGATGCTGTCGCCAAAAAATAAATAAATAA 3106
||||| 1184 AAATAAATGATTACTGTCAAAAAAAAAAAAAAAAAAAAA 1222
Qy 1180 ACACAAATAAATGATTACTGTCAAAAAAAAAAAAAAAAAAAAA 1220

RESULT 37
LOCUS AF090113 5433 bp mRNA ROD 20-SEP-1998
DEFINITION Rattus norvegicus AMPA receptor binding protein mRNA, complete cds.
ACCESSION AF090113
NID 93639076
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 5433)
AUTHORS Srivastava, S., Osten, P., Villim, F.S., Khatri, L., Inman, G.,
Weinberg, R.J. and Ziff, E.B.
Novel anchorage of GluR2/3 to the postsynaptic density by the AMPA
receptor binding protein, ABP
Neuron (1998) In press
2 (bases 1 to 5433)
Direct Submission
AUTHORS Srivastava, S., Villim, F.S., Khatri, L., Daly, C. and Ziff, E.B.
Submitted (03-SEP-1998) Howard Hughes Medical Institute and Dept.
of Biochemistry, New York University, 550, First Avenue, New York,
NY 10016, USA

FEATURES
LOCATION/Qualifiers
SOURCE 1..5433
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
295..2763
/note="contains six PD2 domains"
/codon_start=1
/product="AMPA receptor binding protein"
/db_xref="PID:93639077"
/translation="MVELIKRGSTGLTISGGTDKDKPVSRLNRPGLAARSDLLN
VGDIYRSVNGIRLRLHDEITLLKNVGERVLEVEYELPPAPPENPNRIISKTVDY
SLYKSGNSFGVLRGGAHDLHKSRLVLYTVRPGGPADRGSLKVDGLSIDIGPL
HGASHATAIATLQOCGSHALFOVEYDVATPDVTVANASGPLVVEIAKTPGSGALGSLTT
GSHRNKPAITIDRIKPSVVDVRSALHAGDHILADIGTSTHCSLVEATKLLASVTEK
VRIELPAPOSRRLPKPPEAVRIQRSEQLHNDPCVPSCHSPRPSHCRAFTWAPGGOD

BASE COUNT 1255 a 1067 c 1115 g 1187 t
IGIN

```

		gap of	unknown	length
*	118904	127925:	contig of 9022 bp in length	
*			gap of unknown length	
*	127926	136715:	contig of 8790 bp in length	
*			gap of unknown length	
*	136716	145281:	contig of 8566 bp in length	
*			gap of unknown length	
*	145282	153810:	contig of 8529 bp in length	
*			gap of unknown length	
*	153811	160405:	contig of 6595 bp in length	
*			gap of unknown length	
*	160406	166959:	contig of 6554 bp in length	
*			gap of unknown length	
*	166960	173365:	contig of 6406 bp in length	
*			gap of unknown length	
*	173366	179241:	contig of 5876 bp in length	
*			gap of unknown length	
*	179242	184796:	contig of 5555 bp in length	
*			gap of unknown length	
*	184797	190171:	contig of 5375 bp in length	
*			gap of unknown length	
*	190172	194849:	contig of 4678 bp in length	
*			gap of unknown length	
*	194850	199067:	contig of 4218 bp in length	
*			gap of unknown length	
*	199068	202970:	contig of 3903 bp in length	
*			gap of unknown length	
*	202971	206570:	contig of 3600 bp in length	
*			gap of unknown length	
*	206571	210128:	contig of 3558 bp in length	
*			gap of unknown length	
*	210129	213610:	contig of 3482 bp in length	
*			gap of unknown length	
*	213611	216985:	contig of 3375 bp in length	
*			gap of unknown length	
*	216986	219944:	contig of 2959 bp in length	
*			gap of unknown length	
*	219945	222338:	contig of 2394 bp in length	
*			gap of unknown length	
*	222339	224645:	contig of 2307 bp in length	

Query Match 2.4%; Score 29; DB 28; Length 5433;  
Best Local Similarity 87.2%; Pred. No. 1.76e+00;  
Matches 34; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

★	179242	gap of unknown length
★	184796:	contig of 5555 bp in length
★		gap of unknown length

*	190172	194849:	contig of 4678 bp in length
*			gap of unknown length
*	194850	199067:	contig of 4218 bp in length
*			gap of unknown length
*	199068	202970:	contig of 3903 bp in length
*			gap of unknown length
*	202971	206570:	contig of 3600 bp in length
*			gap of unknown length
*	206571	210128:	contig of 3558 bp in length
*			gap of unknown length
*	210129	213610:	contig of 3482 bp in length
*			gap of unknown length
*	213611	216985:	contig of 3375 bp in length
*			gap of unknown length
*	216986	219944:	contig of 2959 bp in length
*			gap of unknown length
*	219945	222338:	contig of 2394 bp in length
*			gap of unknown length
*	222339	224645:	contig of 2307 bp in length

[illegible]

Contact: Richard M. Myers  
Stanford Human Genome Center (SHGC)  
Stanford University School of Medicine  
Department of Genetics, M-344, Stanford, CA 94305, USA  
Tel: 4157259687  
Fax: 4157259689  
Email: myers@shgc.stanford.edu

Primer A: ACTTTTCTCCTTCAGAGTCACC  
Primer B: TGCACCTTCATTCACATGCA  
STS size: 132  
PCR Profile:

Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds  
Annealing: 62 degrees C for 23 seconds  
Polymerization: 72 degrees C for 30 seconds  
PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9600

Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
Taq Polymerase: 0.05 units/ul  
Total Vol: 10 ul

Buffer: MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 20 mM  
pH: 8.3

Prepared with primer pairs derived from R05470--Merck/UniEST.

#### FEATURES

Source  
STS  
primer\_bind 83..106  
primer\_bind complement(195..214)  
BASE COUNT 72 a 45 c 31 g 79 t 5 others  
ORIGIN

Query Match 2.3%; Score 28; DB 30; Length 232;  
Best Local Similarity 81.0%; Pred. No. 4.65e+00;  
Matches 34; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 1 TTTTCTTTTCTTTTTCAGTNTTCAGTTTATTATGNT 42  
1220 TTTTCTTTTCTTTTTCAGTNTTCAGTTTATTATGNT 1179

RESULT 40  
LOCUS G37172 350 bp DNA STS 30-MAR-1998  
DEFINITION SHGC-57073 Human Homo sapiens STS genomic, sequence tagged site.  
ACCESSION G37172  
NID 92996823  
KEYWORDS STS.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 350)  
AUTHORS Myers,R.M.  
TITLE Human STSs (1997)  
JOURNAL Unpublished (1997)  
COMMENT

Contact: Richard M. Myers  
Stanford Human Genome Center (SHGC)  
Stanford University School of Medicine  
Department of Genetics, M-344, Stanford, CA 94305, USA  
Tel: 4157259687  
Fax: 4157259689  
Email: myers@shgc.stanford.edu

Primer A: CAACCTTGCAGAAATCCTGGAAG  
Primer B: TACCAGGGTGCCAGAAAAG  
STS size: 114  
PCR Profile:

Initial incubation: 95 degrees C for 10 minutes

Denaturation: 94 degrees C for 30 seconds  
Annealing: 60 degrees C for 30 seconds  
Polymerization: 72 degrees C for 23 seconds  
PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9700

Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
Amplifrag Gold Polymerase: 0.07 units/ul  
Total Vol: 5 ul

Buffer: MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 10 mM  
pH: 8.3

Prepared with primer pairs derived from W04555 -- Unigene.

#### FEATURES

Location/Qualifiers  
1..350  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="9"  
/clone\_lib="Human"  
234..347  
primer\_bind 234..255  
primer\_bind complement(328..347)  
BASE COUNT 106 a 56 c 57 g 131 t  
ORIGIN

Query Match 2.3%; Score 28; DB 30; Length 350;  
Best Local Similarity 86.8%; Pred. No. 4.65e+00;  
Matches 33; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 5 TTTTCTTTTCTTTTTCAGTNTTCAGTTTATTATGNT 42  
1222 TTTTCTTTTCTTTTTCAGTNTTCAGTTTATTATGNT 1185

RESULT 41  
LOCUS SCU41441 571 bp DNA INV 02-MAR-1996  
DEFINITION Drosophila melanogaster macrolide binding protein (FKBP12) gene, complete cds.

ACCESSION U41441  
NID 91209721

KEYWORDS fruit fly.

ORGANISM Drosophila melanogaster  
Eukaryota; Eukaryota; Eukaryotes; Metazoa; Arthropoda;  
Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 571)  
AUTHORS Wang,T., Li,B.Y., Danielson,P.D., Shah,P.C., Rockwell,S.,  
Lechleider,R.J., Martin,J., Manganaro,T. and Donahoe,P.K.  
TITLE The Immunophilin FKBP12 Functions as a Common Inhibitor of the  
TGF-beta Family Type I Receptors

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 571)

AUTHORS Wang,T. and Donahoe,P.K.  
TITLE Direct Submission

JOURNAL Submitted (28-NOV-1995) Tongwen Wang, Pediatric Surgery, Mass.  
General Hosp., Fruit St. - WRN1121, Boston, MA 02114, USA

#### FEATURES

Location/Qualifiers  
1..571  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Soares_Placentia_INHP"
/clone="IMAGE:137079"

BASE COUNT      233 a      101 c      110 g      204 t
ORIGIN

Query Match      2.3%; Score 28; DB 26; Length 648;
Best Local Similarity 86.8%; Pred. No. 4.65e+00;
Matches 33; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Ddb 3 TTTTTTTTTTTTTTTTAAACCGTTAAACATTTTTTATT 40
|||||
Cp 1222 TTTTTTTTTTTTTTTTGACAGTAAATCAATTTTATT 1185
|||||

```

RESULT	43
LOCUS	S35960 739 bp mRNA PRI
DEFINITION	laminin receptor homolog [3' region] (human, mRNA Partial, 739 nt).
ACCESSION	S35960
NID	NID 9249370
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 739)
AUTHORS	Bignon,C., Roux-Dosseto,M., Zeigler,M.E., Wicha,M.S. and Martin,P.M.
TITLE	cDNA cloning and genomic analysis of a new multigene family sharing common phylogenetic and expression profiles with the laminin

JOURNAL Biochem. Biophys. Res. Commun. 184 (3), 1165-1172 (1992)  
MEDLINE 92272704  
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 103309] from the original journal article.  
This sequence comes from Fig 2a.

FEATURES  
    source  
        1. .739 /organism="Homo sapiens"  
          /db\_xref="taxon:9606"  
        2. .463  
        partial  
        /note="clone 48-1"  
        /gene="laminin receptor homolog"  
        2. .463  
CDS

```

/partici
/genes= laminin receptor homolog"
/note="C-terminal. This sequence comes from Fig 2a."
/codon_start=1
/db_xref="PID:g249371"
/translation="EDPGVANGRRPARCYRYCKNKPYKSRFCRGVDPDAKIRFDLGLG
KKAKVEFLPCGHMVSDVEYLPOLSEALEAARICANKYMKSCGDKGFHRYRLVLPFFH
IRYNKMLSCAGDRLOLTGMGAGKQALPGFTLAKLSCPSAPSCSTRSN"
183 a 197 c 194 g 165 t
BASE COUNT
ORIGIN

```

Db	697	CACCAATAAATTCCTCTGTCACAAAAA	1738	
Qy	1181	CACAAATAAATTCATTTACTGTCAAAAAA	1222	
RESULT	44			
LOCUS	AF070648	1332 bp	mRNA	PRI
DEFINITION	Homo sapiens clone 24651 mRNA sequence.			
Best Local Similarity 83.3%. Pred. No. 4.65e+00;				
Matches	35;	Conservative	0;	Mismatches 7; Indels 0; Gaps 0;

ACCESSION	AF070066
NID	93283922
KEYWORDS	FLI-CDNA,
SOURCE	human.

Matches 36; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 1388 AAAATAAATAAAAAATTATATAATTTAAAAA 1431

RESULT	48
LOCUS	AF070588      1841 bp      mRNA      PRI      05-AUG-1998
DEFINITION	Homo sapiens clone 24554 unknown mRNA.
ACCESSION	AF070588
NID	G3387961
KEYWORDS	FLI CDNA.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. Anderson,B., Wentland,M.A., Ricafrente,J.Y., Liu,W. and Gibbs,R.A. TITLE A 'double adaptor' method for improved shotgun library construction JOURNAL Anal. Biochem. 236 (1), 107-113 (1996) MEDLINE 95207227 REFERENCE 2 (bases 1 to 1841)
AUTHORS	Yu,W., Anderson,B., Worley,K.C., Muzyk,D.M., Ding,Y., Liu,W., Ricafrente,J.Y., Wentland,M.A., Lennon,G. and Gibbs,R.A.
TITLE	Large-scale concatenation cDNA sequencing
JOURNAL	Genome Res. 7 (4), 353-358 (1997)
MEDLINE	97264341

REFERENCE 3 (bases 1 to 1841)  
 AUTHORS Yu, W. and Gibbs, R. A.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-JUN-1998) Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA  
 FEATURES  
 source location/Qualifiers  
 1..1841 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /sex="female"  
 /dev\_stage="infant"  
 /tissue\_type="brain"  
 /clone="I.M.A.G.E. Consortium clone ID 24554"  
 /clone\_lib="INIB"  
 377..1489  
 /codon\_start=1  
 /product="unknown"  
 /db\_xref="PID:g3387962"  
 /translation="MSRRKQTNPVKHCDSEGEDEETTODEVSSHTSEEDGGVVKVEK  
 ELNTEPQGVNGVNEVHEVTGNLSNPDLLCQPLCDGSRGOLIAHVYQHTAAV  
 VSAKSYMPGVCCGPAISPCSGICRHHLIHSEDORSNCVCGARETSHATFNSEKLPEVL  
 NMESLTIVNEGSPSREGDIAFSPVPYAGILLVCNCAAYRKLLEAQTSPSVRKVAL  
 RRONEPLEVRLQLERETAKTSRRDNTEPEEREVRMRDRREAKRLQRMQETDEQRAK  
 RLQRDEMRRLKRANETPEKQARLIREAREKRLKRLKMKMLRAQFGQDPDSAMAA  
 LAAMNFFQLPVGSVELDSQLCKMAFEEQNSSLH"  
 BASE COUNT 455 a 480 c 537 g 369 t  
 ORIGIN  
 Query Match 2.3%; Score 28; DB 26; Length 1841;  
 Best Local Similarity 79.2%; Pred. No. 4.63e+00;  
 Matches 38; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 Db 1792 TGTCAAGTCAAATAAATTTTATCTTTAAAAAAAAAAAAAAAAA 1839  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1175 TGTGAACACAATAAATTTGATTACTGTCAAAAAAAAAAAAAAAAA 1222  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 RESULT 49  
 LOCUS MMRNASEB4 1908 bp RNA ROD 03-NOV-1993  
 DEFINITION M. musculus seb4 mRNA.  
 ACCESSION X75316  
 NID 6407467  
 KEYWORDS seb4 gene; SEB4D protein; ssDNA binding protein.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;  
 Murinae; Mus.  
 REFERENCE 1 (bases 1 to 1908)  
 AUTHORS Ruehlmann, A., Gupta, A. and Terhorst, C.  
 TITLE A novel murine RRM-type protein and its human homolog  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1908)  
 AUTHORS Ruehlmann, A. C. E. C.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-SEP-1993) A. Ruehlmann Cec, Medizinische Hochschule Hannover (MHH), Institut fuer Molekularbiologie, OE5250, 14 Raum 2630, 30623 Hannover, FRG  
 FEATURES  
 source location/Qualifiers  
 1..1908 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /cell\_line="EL4 murine T cell line"  
 /clone\_lib="lambda gt22 (clontech)"  
 236..1863  
 /gene="seb4"  
 236..949  
 /gene="seb4"  
 /codon\_start=1  
 /product="SEB4"  
 /db\_xref="PID:g407468"  
 /translation="MLIQACPSVFPFRPSAAPSAMHGRSKDTFTFKIFVGGLPYHT  
 DASLRVYFGFGDIEAVYITDROTCKRGYGVFTWADRAAADRACKDPNPIDGRKA  
 DSRVYFGFGDIEAVYITDROTCKRGYGVFTWADRAAADRACKDPNPIDGRKA





AC P13558:  
 DT 01-JAN-1990 (REL. 13, CREATED)  
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)  
 DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)  
 DE MOBILIZATION PROTEIN MBEA.  
 GN MBEA.  
 OS ESCHERICHIA COLI.  
 OG PLASMID COEL.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 OC ENTEROBACTERIACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89364735.  
 RA BOYD A.C., ARCHER J.A.K., SHERRATT D.J.;  
 RL MOL. GEN. GENET. 217:488-498(1989).  
 CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL TO PROMOTE THE SPECIFIC  
 CC TRANSFER OF THE PLASMID IN THE PRESENCE OF CONJUGATIVE PLASMIDS.  
 CC -1- SIMILARITY: HIGH, TO THE MBAA AND MBKA PROTEINS OF E.COLI.  
 DR EMBL; X15873; G43450; -  
 DR PIR; J00390; J00390.  
 KW PLASMID; MOBILITY PROTEIN; CONJUGATION.  
 SQ SEQUENCE 517 AA; 57808 MW; 445FBC4A CRC32;

Query Match 8.3%; Score 93; DB 1; Length 517;  
 Best Local Similarity 32.1%; Pred. No. 2.54e+00;  
 Matches 18; Conservative 13; Mismatches 21; Indels 4; Gaps

DB 431 GERLGIADVAYAT-GORDAERAGHAVESAGALERADRTLEP-VIORELEIRE 484  
 QY 26 GPR-RGSGFEITFDLPPEAAPAGRSASNGHR-KRSRRVLPRVVRKOLPVEE 79  
 | | | | : : : : | | | | : : : : | | | | : : : : |  
 | | | | : : : : | | | | : : : : | | | | : : : : |

RESULT 40  
 ID EXTN\_TOBAC STANDARD; PRT; 620 AA.  
 AC P13983;  
 DT 01-JAN-1990 (REL. 13, CREATED)  
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)  
 DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)  
 DE EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN).  
 GN HRPNT3.  
 OS NICOTIANA TABACUM (COMMON TOBACCO).  
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;  
 OC SOLANALES; SOLANACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. XANTHI; TISSUE-LEAF;  
 RX MEDLINE; 90128263.  
 RA KELLER B., LAMB C.J.;  
 RL GENES DEV. 3:1639-1646(1989).  
 CC -1- FUNCTION: HAS A SPECIALIZED STRUCTURAL FUNCTION, POSSIBLY IN  
 CC THE MECHANICAL PENETRATION OF THE CORTEX AND EPIDERMIS OF THE  
 CC MAIN ROOT.  
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.  
 CC -1- PTM: EXTENSINS CONTAIN A CHARACTERISTIC REPEAT OF THE PENTAPEPTIDE  
 CC SER-PRO(4). THE PROLINE RESIDUE IS HYDROXYLATED AND THEN  
 CC GLYCOSYLATED.  
 CC EMBL; X13885; G19867; -  
 DR PIR; S06733; S06733.  
 DR HSSP; P19989; ICLG.  
 DR REPEAT; CELL WALL; GLYCOPROTEIN; SIGNAL; STRUCTURAL PROTEIN;  
 KW HYDROXYLATION.  
 FT SIGNAL 1 ?  
 FT CHAIN ?  
 FT CHAIN ? 620 EXTENSIN.  
 FT REPEAT 70 73 H-A-P-P.  
 FT REPEAT 148 151 H-A-P-P.  
 FT DOMAIN 229 242 2 X 7 AA TANDEM REPEATS OF T-H-R-H-A-P-P.  
 FT REPEAT 229 235 1.  
 FT REPEAT 236 242 2.  
 FT DOMAIN 205 620 CONTAINS THE SER-PRO(4) REPEATS.  
 FT DOMAIN 499 600 3 X APPROXIMATE TANDEM REPEATS.  
 SQ SEQUENCE 620 AA; 65406 MW; EC549236 CRC32;

Query Match 8.3%; Score 92; DB 1; Length 620;

FT	BINDING	100	100	ATP (BY SIMILARITY).
FT	ACT_SITE	195	195	BY SIMILARITY.
FT	DOMAIN	597	604	POLY-ALA.
FT	VARSPLIC	328	356	MISSING (IN VARIANT 6).
FT	VARSPLIC	378	382	MISSING (IN VARIANT 3).
FT	VARSPLIC	412	448	MISSING (IN VARIANT 3).
FT	VARSPLIC	536	557	AITGVSPRATDPPSHLGGPPA -> GSLTFCFQPRGHVWE
FT				MGMUGV (IN VARIANTS 2, 5, AND 6).
FT	VARSPLIC	558	631	MISSING (IN VARIANTS 3 AND 4).
FT	VARSPLIC	552	631	LDGPPAVAGOCPLVGPQPMHRRLLPARIPRGGLSEARC
FT				LLLFALAAATLCTGLVAVTGLTIPWCFPGATFAP
FT				-> MAPRPKLWASRWGOACPTAVTCCSLPGSLGLAYPRR
FT				VACSCPLLNLPPHWAALGWPIPAVSPQSGVSRPPSP
FT				EP (IN VARIANTS 3 AND 4).
FT	SEQUENCE	631 AA;	69601 MW;	C5E410E8 CRC32;
FT	SEQUENCE	631 AA;	69601 MW;	C5E410E8 CRC32;

Query Match 8.3%; Score 92; DB 1; Length 631;  
Best Local Similarity 35.5%; Pred.No. 3.34e+00;  
Matches 27; Conservative 15; Mismatches 27; Indels 7; Gaps 7

Db	518	HVRQLQERMEMLQAPGAAGAAITGVSPR-ATDPPSH-LDG-P-PAVAGOCPLVGPQPMHR	573
QY	3	HSRCSCHPTWTILQATPAPST-IPGRRCSGPEITFDLPPEAAPACGRPSASRCHRRK	61
Db	574	-RHLLPARIPRGLS	588
QY	62	SRVLYP-RVVRQLP	76

RESULT 42

ID	FIBL_HUMAN	STANDARD;	PRT;	820 AA.
AC	P22105;			
DT	01-AUG-1991 (REL. 19, CREATED)			
DT	01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	FIBRINOGEN-LIKE PROTEIN (FRAGMENT).			
DE	FIBRINOGEN-LIKE PROTEIN (FRAGMENT).			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; PRIMATES.			
RP	SEQUENCE FROM N.A.			
RN	[1]			
RX	MEDLINE: 89367293.			
RL	MOREL Y., BRISTOW J., GITEMAN S.E., MILLER W.L.;			
RL	PROC. NATL. ACAD. SCI. U.S.A. 86:6582-6586(1989).			
CC	- FUNCTION: NOT KNOWN, THIS PROTEIN IS ENCODED ON THE OPPOSITE			
CC	STRAND OF THE STEROID 21-HYDROXYLASE/COMPLEMENT COMPONENT C4			
CC	GENE LOCUS.			
CC	- SIMILARITY: CONTAINS A DOMAIN RELATED TO THE C-TERMINAL PART OF			
CC	THE BETA AND GAMMA CHAINS OF FIBRINOGEN.			
CC	- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III-LIKE DOMAINS.			
DR	EMBL: M25813; G183070; --			
DR	PIR: A33725; A33725.			
DR	HSSP: P24821; ITEN.			
DR	PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; 1.			
KW	REPEAT.			
FT	NON_TER	1		
FT	DOMAIN	22	113	FIBRONECTIN TYPE-III.
FT	DOMAIN	126	221	FIBRONECTIN TYPE-III.
FT	DOMAIN	226	310	FIBRONECTIN TYPE-III.
FT	DOMAIN	328	416	FIBRONECTIN TYPE-III.
FT	DOMAIN	418	508	FIBRONECTIN TYPE-III.
FT	DOMAIN	509	596	FIBRONECTIN TYPE-III.
FT	DOMAIN	597	820	FIBRINOGEN BETA/GAMMA.
FT	DISULFID	606	636	BY SIMILARITY.
FT	DISULFID	758	771	BY SIMILARITY.
FT	SEQUENCE	820 AA;	89601 MW;	EF594B45 CRC32;

Query Match 8.3%; Score 92; DB 1; Length 820;  
Best Local Similarity 33.3%; Pred.No. 3.34e+00;  
Matches 20; Conservative 14; Mismatches 25; Indels 1; Gaps 1

Db	75	QRTVTYEDLEPGKKYKFLYGLGGRKRLGVPVSLGALMTAPEETPP-APELAPEAPEPEEP	133
----	----	--	-----

[illegible]

[illegible]

OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88315020.  
 RA MOHAMED A.H., CHIRALA S.S., MODY N.H., HUANG W.Y., WAKIL S.J.;  
 RL J. BIOL. CHEM. 263:12315-12325(1988).  
 RN [2]  
 RP MUTAGENESIS OF GLY-1257.  
 RC STRAIN-S288C;  
 RX MEDLINE; 94316198.  
 RA OKURA S.;  
 RL MOL. GEN. GENET. 244:90-96(1994).  
 CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF  
 LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.  
 CC THE ALPHA SUBUNIT CONTAINS DOMAINS FOR: ACYL CARRIER PROTEIN,  
 CC 3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE, AND 3-OXOACYL-[ACYL-  
 CC CARRIER-PROTEIN] SYNTHASE. THIS SUBUNIT COORDINATES THE BINDING  
 CC OF THE SIX BETA SUBUNITS TO THE ENZYME COMPLEX.  
 CC -1- CATALYTIC ACTIVITY: ACETYL-COA + N MALONYL-COA + 2N NADPH -  
 CC LONG-CHAIN FATTY ACID + (N+1) COA + N CO(2) + 2N NADP(+).  
 CC -1- CATALYTIC ACTIVITY: ACYL-[ACYL-CARRIER PROTEIN] + MALONYL-[ACYL-  
 CC CARRIER PROTEIN] - 3-OXOACYL-[ACYL-CARRIER PROTEIN] + CO(2) +  
 CC [ACYL-CARRIER PROTEIN].  
 CC -1- CATALYTIC ACTIVITY: (3R)-3-HYDROXYACYL-[ACYL-CARRIER PROTEIN] +  
 CC NADP(+) - 3-OXOACYL-[ACYL-CARRIER PROTEIN] + NADPH.  
 CC -1- SUBUNIT: FATTY ACID SYNTHETASE ARE [ALPHA(6)BETA(6)] HEXAMERS OF  
 CC TWO MULTIFUNCTIONAL SUBUNITS (ALPHA & BETA).  
 CC -1- SIMILARITY: TO THE FATTY ACID SYNTHETASE, SUBUNIT ALPHA FROM  
 CC OTHER FUNGI.  
 DR EMBL; J03936; G171502; -.  
 DR PIR; A31107; A31107.  
 DR SGD; L000602; FAS2.  
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.  
 DR PROSITE; PS00606; B\_KETORACYL SYNTHASE; 1.  
 KW FATTY ACID BIOSYNTHESIS; MULTIFUNCTIONAL ENZYME; OXIDOREDUCTASE;  
 KW TRANSFERASE; NADP; PHOSPHOPANTHETHEINE.  
 FT DOMAIN 1 ? ACYL CARRIER.  
 FT DOMAIN ? ?  
 FT DOMAIN ? 1894 BETA-KETORACYL REDUCTASE.  
 FT BINDING 180 BETA-KETORACYL SYNTHASE.  
 FT ACT\_SITE 1312 1312 PHOSPHOPANTHETHEINE (BY SIMILARITY).  
 FT MUTAGEN 1257 1257 BETA-KETORACYL SYNTHASE (BY SIMILARITY).  
 FT G->S: CERULENIN-RESISTANCE.  
 SQ SEQUENCE 1894 AA; 208098 MW; 310BD492 CRC32;  
 Query Match 8.28; Score 91; DB 1; Length 1894;  
 Best Local Similarity 36.08; Pred. No. 4.37e+00;  
 Matches 18; Conservative 13; Mismatches 18; Indels 1; Gaps 1;  
 104 AKEEAPAPTAAPAAAPAPVAAA-APAAAAAIADEPVKASLLH 152  
 101 AEEGVPALPPEDAPNAASLAPVSPVLEPFLNLTSEPSYALDLSTFLQ 150  
 RESULT 50  
 ID FGF4\_CHICK STANDARD; PRT; 194 AA.  
 AC P48804;  
 DT 01-FEB-1996 (REL. 33, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE FIBROBLAST GROWTH FACTOR-4 PRECURSOR (FGF-4) (HBGF-4).  
 GN FGF4 OR FGF-4.  
 OS GALLUS GALLUS (CHICKEN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;  
 OC GALLIFORMES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 95021713.  
 RA NISWANDER L., JEFFREY S., MARTIN G.R., TICKLE C.;  
 RL NATURE 371:609-612(1994).  
 CC -1- FUNCTION: MAY BE THE ENDOGENOUS RIDGE SIGNAL FOR LIMB DEVELOPMENT.  
 CC -1- RECIPROCAL INTERACTIONS MAY CREATE A POSITIVE FEEDBACK LOOP

CC BETWEEN SONIC HEDGEHOG (SHH) AND FGF4.  
 CC -1- TISSUE SPECIFICITY: POSTERIOR RIDGE.  
 CC -1- INDUCTION: BY RETINOIC ACID.  
 CC -1- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.  
 DR EMBL; U14654; G609348; -.  
 DR PROSITE; PS00247; HBGF\_FGF; 1.  
 KW GROWTH FACTOR; MITOGEN; SIGNAL.  
 FT SIGNAL 1 ? POTENTIAL.  
 FT CHAIN 1 ? FIBROBLAST GROWTH FACTOR-4.  
 SQ SEQUENCE 194 AA; 21642 MW; 5377A763 CRC32;  
 Query Match 8.18; Score 90; DB 1; Length 194;  
 Best Local Similarity 35.68; Pred. No. 5.71e+00;  
 Matches 16; Conservative 14; Mismatches 11; Indels 4; Gaps 4;  
 DB 25 PPFGRLPPGPR-QRRWDA-ALFARSVAR-LPAERRDAARDGYLL 66  
 QY 47 APAGR-PSASRGRKSRRLYPRVVRQLPVEEPNPAKRLFL 90